TITLE: POLYPEPTIDES OF *ALICYCLOBACILLUS SP.*

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims, under 35 U.S.C. 119, priority or the benefit of Danish application Nos. PA 2004 00010, filed January 6, 2004, and PA 2004 00165, filed February 4, 2004, filed, the contents of which are fully incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to functional polypeptides encoded by polynucleotides comprised in the genome of *Alicyclobacillus sp.* deposited under deposit accession number DSM 15716. The invention relates further to the polynucleotides and constructs of such polynucleotides encoding such polypeptides or facilitating their expression as well as to method for preparing the polypeptide. Still further the invention relates to compositions comprising the polypeptide and to uses of the polypeptide.

BACKGROUND OF THE INVENTION

Some enzymes from the genus of *Alicyclobacillus species* are known such as described in Matzke et al.; *Gene cloning, nucleotide sequence and biochemical properties of a cytoplasmic cyclomaltodextrinase (neopullulanase) from Alicyclobacillus acidocaldarius ATCC 2700; reclassification of a group of enzymes*; Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases or Koivula et al.; *Cloning and sequencing of a gene encoding acidophilic amylase from Bacillus acidocaldarius*. J. Gen. Microbiol. 139:2399 (1993) or Bartolucci et al.; *Thioredoxin from Bacillus acidocaldarius: characterization, high-level expression in Escherichia coli and molecular modeling*; Biochem. J. 328:277 (1997) or Tsuruoka et al.; *Collagenolytic Serine-Carboxyl Proteinase from Alicyclobacillus sendainensis Strain NTAP-1: Purification, Characterization, Gene Cloning, and Heterologous Expression;* Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases; Eckert K. & Schneider E., *A thermoacidophilic endoglucanase (celB) from Alicyclobacillus acidocaldarius displays high sequence similarity to arabinofuranosidases belonging to family 51 of glycosyl hydrolases*; Eur. J. Biochem., 270: 3593-3602, 2003.

In the pursuit of novel enzymes it is also known to screen for such new enzymes by subjecting potential candidates to specific enzyme assays. This approach is limited to the availability of enzyme assays and does not allow the identification of functional enzymes or polypeptides for which the activity is still unknown.

Further, whole genome sequencing is a known method to obtain the information on all genes from a given microorganism e.g. as described in Fleischmann et al.; *Whole genome sequences and assembly of Haemophilus influenzae Rd*; Nature 269: 496- 512; (1995).

Most enzymes for industrial use are enzymes which are secreted to the medium by a microorganism. However, only a few percent of a microorganisms' genome encodes secreted proteins. For example only approx. 4% of the *Bacillus subtilis* genome or its closest relatives encode secreted proteins (Van Dijl et al.: *Protein transport pathways in Bacillus subtilis: a genome-based road map*; in "Bacillus subtilis and its closest relatives" - *From genes to cells*; p.337-355; A. L. Sonenshein (ed.); ASM Press 2002).

One disadvantage of genome sequencing is that the vast majority of the obtained sequences encode non secreted proteins.

Also known is signal trapping which is a method to identify genes including nucleotides encoding a signal peptide using a translational fusion to an extra cellular reporter gene lacking its own signal (WO 01/77315).

SUMMARY OF THE INVENTION

The present inventors have found a strain of Alicyclobacillus namely Alicyclobacillus sp. DSM 15716 which grows at low pH (approx 4-5) and at high temperature (50-60 °C). This strain is interesting because the phylogenetic distance between the public known strains and strain DSM 15716 is significant and because the growth conditions are similar to conditions for several applications for industrial enzymes.

The genome of a microorganism contains thousands of different genes; some encoding polypeptides some coding for RNAs. Only a limited number of the genes in the genome of a microorganism encode functional polypeptides which are secreted by the microorganism to the surrounding medium serving an external purpose for the microorganism. Such polypeptides are interesting for industry from the point of view that such polypeptides may be produced in considerable amounts in continuous processes without destroying the cells producing the polypeptides.

It is an object of the present invention to identify and provide polypeptides secreted from *Alicyclobacillus sp.* deposited under deposit accession number DSM 15716 which have functional purpose for the *Alicyclobacillus sp.* because such polypeptides may not only be used for industrial purposes but they may also be produced in industrially relevant processes and amounts.

The present invention provides in a first aspect an isolated polypeptide selected from the group consisting of:

(a) a polypeptide comprising an amino acid sequence which has at least 90% identity with a sequence of a mature polypeptide comprised in the group of SEQ ID NO: 26 to SEQ ID NO:50; and

(b) a polypeptide which is encoded by a nucleotide sequence which hybridize under high stringency conditions with a polynucleotide probe selected from the group consisting of

(i) the complementary strand to a nucleotide sequence selected from the group of regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide.

(ii) the complementary strand to the cDNA sequence contained in a nucleotide sequences selected from the group of regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide

wherein the polypeptide has a function of the corresponding mature polypeptides comprised in SEQ ID NO:26 to SEQ ID NO:50.

In a further aspect the invention provides an isolated enzyme selected from the group consisting of:

(a) an enzyme comprising an amino acid sequence which has at least 90% identity with the amino acid sequence of a mature enzyme selected from the group consisting of acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase secreted from the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716

(b) a polypeptide which is encoded by a nucleotide sequence which hybridize under high stringency conditions with a polynucleotide probe selected from the group consisting of

(i) the complementary strand to a nucleotide sequence comprised in the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716 encoding a mature enzyme selected from the group consisting of acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-

glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase secreted from that strain;

(ii) the complementary strand to the cDNA sequence contained in a nucleotide sequences comprised in the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716 encoding a mature enzyme selected from the group consisting of acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase secreted from that strain

wherein the enzyme have a function selected from acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase

In further aspects the invention provides a polynucleotide encoding the polypeptide of the invention; a nucleotide construct comprising the polynucleotide encoding the polypeptide, operably linked to one or more control sequences that direct the production of the polypeptide in a host cell; a recombinant expression vector comprising the nucleotide construct of the invention and to a recombinant host cell comprising the nucleotide construct of the invention.

In still further aspects the invention provides a method of preparing a polypeptide of the invention comprising:

(a) cultivating a strain comprising a nucleotide sequence encoding a polypeptide of the invention which strain is capable of expressing and secreting the polypeptide and

(b) recovering the polypeptide.

In still further aspects the invention provide a composition comprising a polypeptide of the invention and a method for preparing such a composition comprising admixing the polypeptide of the invention with an excipient.

In still further aspects the invention provides use of the polypeptide of the invention or a composition comprising said polypeptide in various applications.

SEQUENCE LISTING

The present application contains information in the form of a sequence listing, which is appended to the application and also submitted on a data carrier accompanying this application. The contents of the data carrier are fully incorporated herein by reference. The regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide encodes the mature polypeptides of SEQ ID NO:26 to SEQ ID NO:50. The region of SEQ ID NO: 1 encoding a mature polypeptide thus encodes the mature polypeptide sequence comprised in SEQ ID NO:26, the region of SEQ ID NO:2 encoding a mature polypeptide encode the mature polypeptide comprised in SEQ ID NO:27 and so on.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

The term **"identity"** as used herein, is to be understood as the homology between two amino acid sequences or between two nucleotide sequences. For purposes of the present invention, the degree of identity between two amino acid sequences was determined by using AlignX in the program of Vector NTI ver. 7.1 (Informax inc., 7600 Wisconsin Avenue, Suite #1100, Bethesda, MD 20814, USA). Amino acid alignment was created using the Clustal W algorithm (Nucleic Acid Research, 22 (22): 4673-4680, 1994). The following additional parameters are used: Gap opening penalty of 10, Gap extension penalty of 0.05, Gap separation penalty range of 8. Pairwise alignment parameters were Ktuple = 1, gap penalty = 3, gap length opening penalty = 10, gap extension penalty = 0.1, window size = 5 and diagonals = 5. The degree of identity between two nucleotide sequences is determined using the same algorithm and software package as described above for example with the following settings: Gap penalty of 10, and gap length penalty of 10. Pairwise alignment parameters is Ktuple=3, gap penalty=3 and windows=20.

The term **"functional polypeptide"** as used herein in the context of the present invention means a polypeptide which can be expressed and secreted by a cell and which constitutes an operational unit capable of operating in accordance with the function it is designed to fulfil by the

cell. Optionally, co-factors may be required for the polypeptide to adopt the intended function. One example of functional polypeptides is catalytically active polypeptides or enzymes which help the cell catalyzing reactions in the environment surrounding the cell. Another example could be polypeptides which serve as signal substance. Further examples are polypeptides which function as sensors (receptors) for environmental parameters (chemicals in the environment surrounding the cell) or polypeptides, which are active against other organisms (antimicrobial (poly)peptides) or polypeptides, which contributes to the structural integrity of the cell.

The term **"mature region"** as used herein about portion of an amino acid sequences or polypeptide means the portion or region or domain or section of the amino acid sequences or polypeptide which is the mature functional polypeptide.

The term **"region of nucleotide sequence encoding a mature polypeptide"** as used herein means the region of a nucleotide sequence counting from the triplet encoding the first amino acid of a mature polypeptide to the last triplet encoding the last amino acid of a mature polypeptide.

Polypeptides of the invention

The polypeptides of the invention are polypeptides secreted by *Alicyclobacillus sp.* DSM 15716 with the purpose of serving a function for that particular cell and related polypeptides thereof.

Among the thousands of potential genes in the genome of *Alicyclobacillus sp.* DSM 15716 the polynucleotides of this genome encoded 25 secreted functional mature polypeptides comprised in SEQ ID NO: 26 to SEQ ID NO:50, which were determined to be functional, that is translated into functional polypeptides by the chosen host cell.

Accordingly, *Alicyclobacillus sp.* DSM 15716 expresses and secretes the functional mature polypeptides comprised in SEQ ID NO: 26 to SEQ NO: 50 and in the genome of that particular strain, the regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide are the genes encoding the mature polypeptides comprised in SEQ ID NO: 26 to SEQ NO: 50. Further in a particular embodiment the genes encoding the mature polypeptides comprised in of SEQ ID NO:26 to SEQ NO: 50 can all be expressed and their corresponding mature polypeptides can be secreted when culturing an E. coli host transformed with polynucleotides comprising those regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide. By comparing homology or identity of the sequences of the 25 polypeptide sequences to known se-

quences the particular function of the polypeptides were annotated. At least 15 of the 25 secreted functional polypeptides were determined to be enzymes.

The invention provides an isolated polypeptide selected from the group consisting of:

(a) a polypeptide having an amino acid sequence which has at least 90% identity with an amino acid sequence selected from the group consisting of the mature polypeptides comprised in SEQ ID NO: 26 to SEQ ID NO:50 and

(b) a polypeptide which is encoded by a nucleotide sequence which hybridize under high stringency conditions with a polynucleotide probe selected from the group consisting of

(i) the complementary strand to a nucleotide sequence selected from the group consisting of regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide,

(ii) the complementary strand to the cDNA sequence contained in a nucleotide sequences selected from regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide;

wherein the polypeptide exhibits the function of the corresponding mature polypeptide of SEQ ID NO: 26 to SEQ ID NO:50.

In one particular embodiment the polypeptide of the invention is selected among the enzymes secreted by *Alicyclobacillus sp.* deposited under DSM accession No. 15716 and isolated by the present inventors, i.e. the group of enzymes consisting of acid endoglucanase, acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease, HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase, phytase, phospholipase C, polysaccharide deacetylase, xylan deacetylase and sulfite oxidase.

The invention also provides an isolated enzyme selected from the group consisting of:

(a) an enzyme comprising an amino acid sequence which has at least 90% identity with the amino acid sequence of a mature enzyme selected from the group consisting of acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-

prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase secreted from the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716 and

(b) a polypeptide which is encoded by a nucleotide sequence which hybridize under high stringency conditions with a polynucleotide probe selected from the group consisting of

(i) the complementary strand to a nucleotide sequence comprised in the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716 encoding a mature enzyme selected from the group consisting of acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase secreted from that strain;

(ii) the complementary strand to the cDNA sequence contained in a nucleotide sequences comprised in the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716 encoding a mature enzyme selected from the group consisting of acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase secreted from that strain and

wherein the enzyme have a function selected from acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase.

In a particular embodiment the enzyme is an isolated enzyme selected from the group consisting of:

(a) an enzyme having an amino acid sequence which has at least 90% identity with an amino acid sequence selected from mature enzymes comprised in SEQ ID NO: 26 to SEQ ID NO:40 and

(b) an enzyme which is encoded by a nucleotide sequence which hybridize under high stringency conditions with a polynucleotide probe selected from the group consisting of

(i) the complementary strand to a nucleotide sequence selected from the group of regions of SEQ ID NO: 1 to SEQ ID NO: 15 encoding the mature enzyme,

(ii) the complementary strand to the cDNA sequence contained in a nucleotide sequences selected from regions of SEQ ID NO: 1 to SEQ ID NO: 15 encoding the mature polypeptide and

wherein the enzyme has a function of the corresponding mature polypeptides comprised in SEQ ID NO:26 to SEQ ID NO:40

The polypeptide of the invention is an isolated polypeptide, preferably the preparation of the polypeptide of the invention contains at the most 90% by weight of other polypeptide material with which it may be natively associated (lower percentages of other polypeptide material are preferred, e.g. at the most 80% by weight, at the most 60% by weight, at the most 50% by weight, at the most 40% at the most 30% by weight, at the most 20% by weight, at the most 10% by weight, at the most 9% by weight ,at the most 8% by weight, at the most 6% by weight, at the most 5% by weight, at the most 4% at the most 3% by weight, at the most 2% by weight, at the most 1% by weight and at the most ½% by weight). Thus, it is preferred that the isolated polypeptide of the invention is at least 92% pure, i.e. that the polypeptide of the invention constitutes at least 92% by weight of the total polypeptide material present in the preparation, and higher percentages are preferred such as at least 94% pure, at least 95% pure, at least 96% pure, at least 96% pure, at least 97% pure, at least 98% pure, at least 99%, and at the most 99.5% pure. In particular, it is preferred that the polypeptide of the invention is in "essentially pure form", i.e. that the polypeptide preparation is essentially free of other polypeptide material with which it is natively associated. This can be accomplished, for example, by preparing the polypeptide of the invention by means of well-known recombinant methods.

The polypeptide of the invention of the invention may be synthetically made, naturally occurring or a combination thereof. In a particular embodiment the polypeptide of the invention may be obtained from a microorganism such as a prokaryotic cell, an archaeal cell or a eukaryotic cell. The cell may further have been modified by genetic engineering

In a particular embodiment, the polypeptide of the invention is an enzyme exhibiting optimum enzyme activity at a temperature within the range from about 10°C to about 80 °C, particularly in the range from about 20°C to about 60°C.

In a particular embodiment the polypeptide of the invention is an enzyme, which is functionally stabile at a temperature of up to 100 °C, in particular up to 80 °C, more particularly up to 60°C.

In a particular embodiment the polypeptide of the invention is an enzyme exhibiting at least 20%, in particular at least 40%, such as at least 50%, in particular at least 60%, such as at least 70%, more particularly at least 80%, such as at least 90%, most particularly at least 95%, such as about or at least 100% of the enzyme activity of an enzyme selected from mature enzymes comprised in SEQ ID NO: 26 to SEQ ID NO: 50.

In a particular embodiment the polypeptide of the invention comprises, contains or consists of an amino acid sequence which has at least 90% identity with a polypeptide sequence selected from the group consisting of mature polypeptides comprised in SEQ ID NO: 26 to SEQ ID NO: 50; particularly at least 95%, e.g. at least 96%, such as at least 97%, and even more particularly at least 98%, such as at least 99% or even 100% identity.

In another particular embodiment the polypeptide of the invention comprises, contains or consists of an amino acid sequence, which has at least 50% identity with a polypeptide sequence selected from the group consisting of mature polypeptides comprised in SEQ ID NO: 26 to SEQ ID NO: 50; particularly at least 60%, particularly at least 65%, particularly at least 70%, particularly at least 75%, particularly at least 80%, and even more particularly at least 85% identity.

In a particular embodiment, the amino acid sequence of the polypeptide of the invention differs by at the most ten amino acids (e.g. by ten amino acids), in particular by at the most five amino acids (e.g. by five amino acids), such as by at the most four amino acids (e.g. by four amino acids), e.g. by at the most three amino acids (e.g. by three amino acids), in particular by at the most two amino acids (e.g. by two amino acids), such as by one amino acid from the mature polypeptides comprised in SEQ ID NO: 26 to SEQ ID NO: 50.

The polypeptide of the invention may be a wild-type polypeptide isolated from a natural source such as the strain *Alicyclobacillus sp.* DSM 15716 or another wild type strain, however the present invention also encompass artificial variants, where a polypeptide of the invention has been mutated for example by adding, substituting and/or deleting one or more amino acids from said polypeptide while retaining the function of the polypeptide and/or other properties. Hence, the polypeptide of the invention may be an artificial variant, wherein at least one substi-

tution, deletion and/or insertion of an amino acid has been made to an amino acid sequence comprising or consisting of the mature polypeptide comprised in SEQ ID NO: 26 to SEQ ID NO: 50.

The polypeptides of the invention also include functional fragments of the amino acid sequences described herein and nucleic acids encoding functional fragments of the amino acid sequences described herein, including fragments of the mature enzymes secreted from the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, as described herein, including fragment of an enzyme selected from the group consisting of acid endoglucanase, acid cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl protease, serine protease, HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase, phytase, phospholipase C, polysaccharide deacetylase, xylan deacetylase and sulfite oxidase secreted from the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716.

Artificial variants may be constructed by standard techniques known in the art usually followed by screening and/or characterization. Standard techniques includes classical mutagenesis, e.g. by UV irradiation of the cells or treatment of cells with chemical mutagens as described by Gerhardt et al. (1994); in vivo gene shuffling as described in WO 97/07205; in vitro shuffling as described by Stemmer, (1994) or WO 95/17413, random mutagenesis as described by Eisenstadt E. et al., (1994); PCR techniques as described by Poulsen et al. (1991); family shuffling as described by J.E. Ness, et al, Nature Biotechnology, vol. 17, pp. 893-896 (1999); site-directed mutagenesis as described by Sambrook et al. (1989), Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor, NY. A general description of nucleotide substitution can be found in e.g. Ford et al., 1991, Protein Expression and Purification 2, p. 95-107.

Such standard genetic engineering methods may also be used prepare a diversified library of variant nucleotide sequences from the genes encoding one or more parent enzymes of the invention, expressing the enzyme variants in a suitable host cell and selecting a preferred variant(s). A diversified library can be established by a range of techniques known to the art (Reetz MT; Jaeger KE, in Biocatalysis - from Discovery to Application edited by Fessner WD, Vol. 200, pp. 31-57 (1999); Stemmer, Nature, vol. 370, p.389-391, 1994; Zhao and Arnold, Proc. Natl. Acad. Sci., USA, vol. 94, pp. 7997-8000, 1997; or Yano et al., Proc. Natl. Acad. Sci., USA, vol. 95, pp 5511-5515, 1998).

In a particular embodiment of the invention, amino acid changes (in the artificial variant as well as in wild-type enzyme) are of a minor nature, that is conservative amino acid substitutions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 20-25 residues; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

Examples of conservative substitutions are within the group of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine, valine and methionine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine and threonine). Amino acid substitutions which do not generally alter and or impair the function of a protein are known in the art and are described, for example, by H. Neurath and R.L. Hill, 1979, In, The Proteins, Academic Press, New York. The most commonly occurring exchanges are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, and Asp/Gly as well as these in reverse.

In a particular embodiment the amino acid changes are of such a nature that the physico-chemical properties of the polypeptides are altered. For example, amino acid changes may be performed, which improve the thermal stability of the enzyme, which alter the substrate specificity, which changes the pH optimum, and the like.

Particularly, the number of such substitutions, deletions and/or insertions in the polypeptide of the invention, particularly in those polypeptides selected from the group consisting of mature polypeptides comprised in SEQ ID NO: 26 to SEQ ID NO: 50 to produce an artificial variant is at the most 10, such as at the most 9, e.g. at the most 8, more preferably at the most 7, e.g. at the most 6, such as at the most 5, most preferably at the most 4, e.g. at the most 3, such as at the most 2, in particular at the most 1.

In a particular embodiment the artificial variant is a variant, which has an altered, preferably reduced, immunogenicity, especially allergenicity, in animals including man as compared to a parent enzyme. The term "immunogenicity" in this context is to be understood as the artificial variant capability of invoking a an altered, in particular reduced, immunological response when administered to an animal, including intravenous, cutaneous, subcutaneous, oral and intratracheal administration. The term "immunological response" in this context means that the administration of the artificial variant causes an alteration in the immunoglobulin levels in the

animal body, such as in IgE, IgG and IgM or an alteration in the cytokine level in the animal body. Methods for mapping immunogenic/antigenic epitopes of a protein, preparing variants with altered immunogenicity and methods for measuring an immunological response is well known to the art and are described e.g. in WO 92/10755, WO 00/26230, WO 00/26354 and WO 01/31989. The term "allergenicity" in this context is to be understood as the artificial variant ability of invoking an altered, in particular reduced, production of IgE in an animal as well as the ability to bind IgE from said animal. Particularly allergenicity arising from intratracheal administration of the polypeptide variant to the animal is particularly of interest (also known as respiratory allergenicity).

In a further embodiment, the polypeptide of the invention is a polypeptide which is encoded by nucleotide sequences which hybridize under at least high stringency conditions, particularly under very high stringency conditions with a polynucleotide probe selected from the group consisting of

(i) the complementary strand to a nucleotide sequence selected from the group of regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide,

(ii) the complementary strand to the cDNA sequence contained in a nucleotide sequences selected from regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide

(iii) a fragment of (i) or (ii) encoding a secreted polypeptide having the function of the corresponding mature polypeptide comprised in SEQ ID NO: 26 to SEQ ID NO: 50

(J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York).

In particular, the polypeptide of the invention is encoded by a polynucleotide comprising a nucleotide sequence selected from the group of regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide or a sequences differing there from by virtue of the degeneracy of the genetic code. More particularly, the polypeptide of the invention is encoded by a polynucleotide consisting of a nucleotide sequence selected from the group of regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide or a sequence differing there from by virtue of the degeneracy of the genetic code.

The nucleotide sequences of regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide or a subsequence thereof, as well as the amino acid sequences of the mature polypeptides comprised in SEQ ID NO: 26 to SEQ ID NO: 50 or a fragment thereof, may be used to design a polynucleotide probe to identify and clone DNA encoding enzymes of the in-

vention from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic or cDNA of the genus or species of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 15, preferably at least 25, more preferably at least 35 nucleotides in length, such as at least 70 nucleotides in length. It is; however, preferred that the polynucleotide probe is at least 100 nucleotides in length. For example, the polynucleotide probe may be at least 200 nucleotides in length, at least 300 nucleotides in length, at least 400 nucleotides in length or at least 500 nucleotides in length. Even longer probes may be used, e.g., polynucleotide probes which are at least 600 nucleotides in length, at least 700 nucleotides in length, at least 800 nucleotides in length, or at least 900 nucleotides in length. Both DNA and RNA probes can be used. The probes are typically labelled for detecting the corresponding gene (for example, with ^{32}P, ^{3}H, ^{35}S, biotin, or avidin).

Thus, a genomic DNA or cDNA library prepared from such other organisms may be screened for DNA, which hybridizes with the probes described above and which encodes enzymes of the invention. Genomic or other DNA from such other organisms may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to, and immobilized, on nitrocellulose or other suitable carrier materials. In order to identify a clone or DNA which has the required homology and/or identity or is homologous and/or identical with of nucleotides selected from regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide, the carrier material with the immobilized DNA is used in a Southern blot.

For purposes of the present invention, hybridization indicates that the nucleotide sequence hybridizes to a labelled polynucleotide probe which again hybridizes to a nucleotide sequence selected from regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide under high to very high stringency conditions. Molecules to which the polynucleotide probe hybridizes under these conditions may be detected using X-ray film or by any other method known in the art. Whenever the term "polynucleotide probe" is used in the present context, it is to be understood that such a probe contains at least 15 nucleotides.

In an interesting embodiment, the polynucleotide probe is the complementary strand of a nucleotide sequence selected from regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide.

In another interesting embodiment, the polynucleotide probe is the complementary strand of a nucleotide sequence which encodes an enzyme selected from the group of SEQ ID

NO: 26 to SEQ ID NO: 50. In a further interesting embodiment, the polynucleotide probe is the complementary strand of a mature polypeptide coding region of a nucleotide sequence selected from regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide.

For long probes of at least 100 nucleotides in length, high to very high stringency conditions are defined as pre-hybridization and hybridization at 42°C in 5X SSPE, 1.0% SDS, 5X Denhardt's solution, 100 microgram/ml sheared and denatured salmon sperm DNA, following standard Southern blotting procedures. Preferably, the long probes of at least 100 nucleotides do not contain more than 1000 nucleotides. For long probes of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 0.1 x SSC, 0.1% SDS at 60°C (high stringency), in particular washed three times each for 15 minutes using 0.1 x SSC, 0.1% SDS at 68°C (very high stringency).

Although not particularly preferred, it is contemplated that shorter probes, e.g. probes which are from about 15 to 99 nucleotides in length, such as from about 15 to about 70 nucleotides in length, may be also be used. For such short probes, stringency conditions are defined as pre-hybridization, hybridization, and washing post-hybridization at 5°C to 10°C below the calculated Tm using the calculation according to Bolton and McCarthy (1962, Proceedings of the National Academy of Sciences USA 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1X Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2 mg of yeast RNA per ml following standard Southern blotting procedures.

For short probes which are about 15 nucleotides to 99 nucleotides in length, the carrier material is washed once in 6X SCC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6X SSC at 5°C to 10°C below the calculated Tm.

SEQ ID NO: 26 acid endoglucanase or acid cellulase

In a particular embodiment the polypeptide of the invention is an acid endoglucanase or acid cellulase comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with an acid endoglucanase or acid cellulase obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature acid endoglucanase or acid cellulase comprised in SEQ ID NO: 26. More specifically the mature acid endoglucanase or acid cellulase comprise or consists of the sequences from position 25 to

959 of SEQ ID NO: 26. In the present context an acid endoglucanase is defined as enzyme, which endohydrolyzes 1,4-beta-D-glucosidic linkages in cellulose, lichenin or cereal beta-D-glucans particularly at acidic conditions. In the present context an acid cellulase is defined as enzyme, which endohydrolyzes 1,4-beta-D-glucosidic linkages in cellulose, particularly at acidic conditions.

SEQ ID NO: 27 aspartyl protease

In a particular embodiment the polypeptide of the invention is an aspartyl protease comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with an aspartyl protease obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature aspartyl protease comprised in SEQ ID NO: 27. More specifically the mature aspartyl protease comprises or consists of the sequences from position 33 to 272 of SEQ ID NO: 27. In the present context an aspartyl protease is defined as defined as an enzyme, which hydrolyses proteins or peptides, and which contains two aspartyl residues in the catalytic site.

Surprisingly, the aspartyl protease comprised in SEQ ID NO: 27 is the first isolated aspartyl protease, in particular family A4 protease, isolated from a bacterium. Furthermore, surprisingly this aspartyl protease differs from known fungal aspartyl proteases by the absence of disulphide bridges in the molecule. The aspartyl protease comprised in SEQ ID NO: 27 contains only one Cystein and thus no disulphide bridges in the protease structure as compared to e.g. SEQ ID NO: 55 disclosing a known fungal aspartyl protease, which are composed of two peptides cross linked by 2 disulphide bridges. Hence, the aspartyl protease of *Alicyclobacillus sp.* specifically that deposited under DSM accession No. 15716 a second propeptide is missing and thus requires one less maturation step less in its production. This is an advantage for the cellular production. Family A4 proteases are known to the art as proteases having an Asp/Asp Asp/Glu configuration in the active site.

SEQ ID NO: 28 or SEQ ID NO:35 Multi copper oxidase

In a particular embodiment the polypeptide of the invention is a multi copper oxidase comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more

particularly at least 99% or most particularly 100% identity with multi copper oxidase obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature multi copper oxidase comprised in SEQ ID NO: 28 or 35. More specifically the mature multi copper oxidase comprises or consists of the sequences from position 26 to 315 of SEQ ID NO: 28 or position 50 to 597 of SEQ ID NO: 35. In the present context a multi-Cu-oxidase is defined as a protein, which possesses at least three spectroscopically different copper centers. Multicopper oxidases can be laccases that oxidizes many different types of phenols and diamines, ascorbate oxidases, ceruloplasmin, that oxidizes a great variety of inorganic and organic substances or part of proteins that have lost the ability to bind copper and thereby mediate heavy metal resistance by sequestration of the heavy metal in the periplasm of the bacterium.

SEQ ID NO: 29 or SEQ ID NO: 30 Serine-carboxyl protease

In a particular embodiment the enzyme of the invention is a serine-carboxyl protease comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with the serine-carboxyl protease obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature serine-carboxyl protease comprised in SEQ ID NO: 29 or 30. More specifically the mature serine-carboxyl protease comprises or consists of the sequences from position 190 to 626 of SEQ ID NO: 29 or position 25 to 533 of SEQ ID NO: 30. In the present context a serine-carboxyl protease is defined as a protease belonging to the Enzyme class EC 3.4.21.100 (pseudomonapepsin) which proteolytic enzymes fold resembles that of subtilisin, with a unique catalytic triad, Ser-Glu-Asp, as well as the presence of an aspartic acid residue in the oxyanion hole. A polypeptide sequence can be classified as a serine-carboxyl peptidase, if the amino acids of the catalytic site are present in the sequence and if it shows peptide sequence similarity to peptide sequences in MEROPS serine protease family 53.

SEQ ID NO: 31 serine protease or a HtrA-like serine protease

In a particular embodiment the polypeptide of the invention is a serine protease or a HtrA-like serine protease comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with

the serine protease or the HtrA-like serine protease obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature serine protease carboxyl protease comprised in SEQ ID NO: 31. More specifically the mature serine protease comprises or consists of the sequences from position 42 to 411 of SEQ ID NO: 31. In the present context a serine protease is defined as an enzyme, which hydrolyses proteins or peptides, and which contains a serine residue in the catalytic site. A HtrA-like protease is defined as an enzyme that degrades damaged proteins in the extra cellular compartment of a bacterial cell at elevated temperatures.

SEQ ID NO: 32 disulfide isomerase

In a particular embodiment the polypeptide of the invention is a disulfide isomerase comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with the disulfide isomerase obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature disulfide isomerase comprised in SEQ ID NO: 32. More specifically the mature disulfide isomerase comprises or consists of the sequences from position 31 to 212 of SEQ ID NO: 32. In the present context a disulphide isomerase is defined as enzyme, which catalyses the rearrangement of both intrachain and interchain disulfide bonds in proteins to form the native structures.

SEQ ID NO:33 gamma-D-glutamyl-L-diamino acid endopeptidase

In a particular embodiment the polypeptide of the invention is a gamma-D-glutamyl-L-diamino acid endopeptidase comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with the gamma-D-glutamyl-L-diamino acid endopeptidase obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature gamma-D-glutamyl-L-diamino acid endopeptidase comprised in SEQ ID NO: 33. More specifically the mature gamma-D-glutamyl-L-diamino acid endopeptidase comprises or consists of the sequences from position 30 to 266 of SEQ ID NO: 33. In the present context a gamma-D-glutamyl-L-diamino acid endopeptidase is defined as an enzyme that hydrolyses gamma-D-glutamyl bonds to (L) meso-diaminopimelic acid in L-Ala-gamma-D-Glu-|-

(L)meso-diaminopimelic acid-(L)-D-Ala. It is required that the omega-amino and omega-carboxyl groups of the (L) meso- diaminopimelic acid group are unsubstituted.

SEQ ID NO: 34 endo-beta-N-acetylglucosaminidase

In a particular embodiment the polypeptide of the invention is an endo-beta-N-acetylglucosaminidase comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with the endo-beta-N-acetylglucosaminidase obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature endo-beta-N-acetylglucosaminidase comprised in SEQ ID NO: 34. More specifically the mature endo-beta-N-acetylglucosaminidase comprises or consists of the sequences from position 27 to 768 of SEQ ID NO: 34. In the present context an endo-beta-N-Acetylglucosaminidase is defined as enzyme that hydrolyses the 1, 4-beta-linkages between N-acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan heteropolymers of the prokaryotes cell walls.

SEQ ID NO:36.peptidyl-prolyl-isomerase

In a particular embodiment the polypeptide of the invention is a peptidyl-prolyl-isomerase comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with the peptidyl-prolyl-isomerase obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature peptidyl-prolyl-isomerase comprised in SEQ ID NO: 36. More specifically the mature peptidyl-prolyl-isomerase comprises or consists of the sequences from position 30 to 246 of SEQ ID NO: 36. In the present context a peptidyl-prolyl-isomerase is defined as an enzyme that accelerates protein folding by catalyzing the cis-trans isomerization of proline imidic peptide bonds in oligopeptides.

SEQ ID NO: 37 acid phosphatase or a phytase or a phospholipase C

In a particular embodiment the polypeptide of the invention is an acid phosphatase or a phytase or a phospholipase C comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with the acid phosphatase or phytase or phospholipase C obtainable from Alicyclobacillus sp., in

particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature acid phosphatase or phytase or phospholipase C comprised in SEQ ID NO: 37. More specifically the mature acid phosphatase or a phytase or a phospholipase C comprises or consists of the sequences from position 28 to 608 of SEQ ID NO: 37. An acid phosphatase is defined as enzyme hydrolyzing an orthophosphoric monoester into an alcohol and phosphate. In the present context a phytase is defined as an enzyme removing a phosphate group from phytate. A phospholipase C is defined as an enzyme hydrolyzing phosphatidylcholine into 1,2-diacylglycerol and choline.

SEQ ID NO: 38 or SEQ ID NO: 39 polysaccharide deacetylase

In a particular embodiment the polypeptide of the invention is a polysaccharide deacetylase or a xylan deacetylase comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with the polysaccharide deacetylase or the xylan deacetylase obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature polysaccharide deacetylase or a xylan deacetylase comprised in SEQ ID NO: 38 or 39. More specifically the mature polysaccharide deacetylase or a xylan deacetylase comprises or consists of the sequences from position 26 to 251 of SEQ ID NO: 38 or position 22 to 324 of SEQ ID NO: 39. In the present context a polysaccharide deacetylase is defined as an enzyme, which removes acetyl residues from a specific acetylated polysaccharide by hydrolysis. A xylan deacetylase is defined as an enzyme removing acetyl groups from acetylated xylan.

SEQ ID NO:40 sulfite oxidase

In a particular embodiment the polypeptide of the invention is a sulfite oxidase comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with the sulfite oxidase obtainable from Alicyclobacillus sp., in particular that strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716, more particularly the mature sulfite oxidase comprised in SEQ ID NO: 40. More specifically the mature sulfite oxidase comprises or consists of the sequences from position 30 to 214 of SEQ ID NO: 40. A sulfite oxidase is defined as enzyme that oxidizes sulfite to sulfate.

SEQ ID NO:41 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO:41. In particular with the mature functional polypeptide comprised in SEQ ID NO: 41. More specifically the mature functional polypeptide comprises or consists of the sequences from position 22 to 257 of SEQ ID NO: 41.

SEQ ID NO:42 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO:42. In particular with the mature functional polypeptide comprised in SEQ ID NO: 42. More specifically the mature functional polypeptide comprises or consists of the sequences from position 25 to 1130 of SEQ ID NO: 42.

SEQ ID NO:43 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO:43. In particular with the mature functional polypeptide comprised in SEQ ID NO: 43. More specifically the mature functional polypeptide comprises or consists of the sequences from position 42 to 248 of SEQ ID NO: 43.

SEQ ID NO:44 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO:44. In particular with the mature functional polypeptide comprised in SEQ ID NO: 44. More specifically the

mature functional polypeptide comprises or consists of the sequences from position 26 to 172 of SEQ ID NO: 44.

SEQ ID NO:45 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO:45. In particular with the mature functional polypeptide comprised in SEQ ID NO: 45. More specifically the mature functional polypeptide comprises or consists of the sequences from position 31 to 242 of SEQ ID NO: 45.

SEQ ID NO:46 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO:46. In particular with the mature functional polypeptide comprised in SEQ ID NO: 46. More specifically the mature functional polypeptide comprises or consists of the sequences from position 25 to 280 of SEQ ID NO: 46.

SEQ ID NO:47 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO:47. In particular with the mature functional polypeptide comprised in SEQ ID NO: 47. More specifically the mature functional polypeptide comprises or consists of the sequences from position 26 to 478 of SEQ ID NO: 47.

SEQ ID NO:48 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO: 48. In particu-

lar with the mature functional polypeptide comprised in SEQ ID NO: 48. More specifically the mature functional polypeptide comprises or consists of the sequences from position 20 to 340 of SEQ ID NO: 48.

SEQ ID NO:49 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO: 49. In particular with the mature functional polypeptide comprised in SEQ ID NO: 49. More specifically the mature functional polypeptide comprises or consists of the sequences from position 30 to 341 of SEQ ID NO: 49.

SEQ ID NO: 50 functional polypeptide

In a particular embodiment the polypeptide of the invention is a functional polypeptide comprising or consisting of an amino acid sequence which has at least 90%, particularly at least 95%, more particularly at least 96%, more particularly at least 97%, more particularly at least 98%, more particularly at least 99% or most particularly 100% identity with SEQ ID NO: 50. In particular with the mature functional polypeptide comprised in SEQ ID NO: 50. More specifically the mature functional polypeptide comprises or consists of the sequences from position 29 to 400 of SEQ ID NO: 50.

Polynucleotides

The present invention also relates to polynucleotides comprising or consisting of a nucleotide sequence encoding a polypeptide of the invention. In a particular embodiment, the nucleotide sequence is set forth in SEQ ID NO: 1 to SEQ ID NO: 25 including nucleotide sequences differing there from by virtue of the degeneracy of the genetic code. In a further embodiment the polynucleotide of the invention is a modified nucleotide sequence which comprises or consists of a nucleotide sequence selected from the regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide and which comprises at least one modification/mutation compared with the parent nucleotide sequence comprised in SEQ ID NO: 1 to SEQ ID NO: 25.

The techniques used to isolate and/or clone a nucleotide sequence encoding an enzyme are known in the art and include isolation from genomic DNA, preparation from cDNA, or a combination thereof. The cloning of the nucleotide sequences of the present invention from

such genomic DNA can be effected, *e.g.*, by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, *e.g.*, Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nucleotide sequence-based amplification (NASBA) may be used.

The nucleotide sequence may be obtained by standard cloning procedures used in genetic engineering to relocate the nucleotide sequence from its natural location to a different site where it will be reproduced. The cloning procedures may involve excision and isolation of a desired fragment comprising the nucleotide sequence encoding the polypeptide, insertion of the fragment into a vector molecule, and incorporation of the recombinant vector into a host cell where multiple copies or clones of the nucleotide sequence will be replicated. The nucleotide sequence may be of genomic, cDNA, RNA, semi-synthetic, synthetic origin, or any combinations thereof.

In particular the polynucleotide comprises, preferably consists of, a nucleotide sequence which has at least 50% identity with a nucleotide sequence selected from the regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide. Particularly, the nucleotide sequence has at least 65% identity, more particularly at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with a nucleotide sequence selected from the regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide. Particularly, the nucleotide sequence comprises a nucleotide sequence selected from the regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide. In an even more particular embodiment, the nucleotide sequence consists of a nucleotide sequence selected from the regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide.

In particular the polynucleotide comprises, preferably consists of, a nucleotide sequence which has at least 50% identity, particularly at least 65% identity, more particularly at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with a nucleotide sequence encoding a mature enzyme selected from acid endoglucanase or acid cellulase, aspartyl protease, multi copper oxidase, serine-

carboxyl protease, serine protease or HtrA-like serine protease, disulfide isomerase, gamma-D-glutamyl-L-diamino acid endopeptidase, endo-beta-N-acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase or phytase or phospholipase C, polysaccharide deacetylase or xylan deacetylase and sulfite oxidase secreted from the strain of *Alicyclobacillus sp.* Deposited under DSM accession No. 15716

SEQ ID NO: 1

In a particular embodiment the polynucleotide of the invention encodes an acid endoglucanase or acid cellulase and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 73 to 2877of SEQ ID NO: 1

SEQ ID NO: 2

In a particular embodiment the polynucleotide of the invention encodes an aspartyl protease and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 97 to 816 of SEQ ID NO: 2

SEQ ID NO: 3 and 10

In a particular embodiment the polynucleotide of the invention encodes an multi copper oxidase and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 76 to 945 of SEQ ID NO: 1 or 148 to 1791 of SEQ ID NO: 10

SEQ ID NO: 4 and 5

In a particular embodiment the polynucleotide of the invention encodes a serine-carboxyl protease and comprises or consists of an nucleotide sequence which has at least 70% identity, more

particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 568 to 1878 of SEQ ID NO: 4 or 73 to 1599 of SEQ ID NO: 5.

SEQ ID NO: 6

In a particular embodiment the polynucleotide of the invention encodes a serine protease or a HtrA-like serine protease and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 124 to 1233 of SEQ ID NO: 6.

SEQ ID NO: 7

In a particular embodiment the polynucleotide of the invention encodes a disulfide isomerase and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 91 to 633 of SEQ ID NO: 7.

SEQ ID NO: 8

In a particular embodiment the polynucleotide of the invention encodes a gamma-D-glutamyl-L-diamino acid endopeptidase and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 88 to 798 of SEQ ID NO: 8.

SEQ ID NO: 9

In a particular embodiment the polynucleotide of the invention encodes a endo-beta-N-acetylglucosaminidase and comprises or consists of an nucleotide sequence which has at least

70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 79 to 2304 of SEQ ID NO: 9.

SEQ ID NO: 11

In a particular embodiment the polynucleotide of the invention encodes a peptidyl-prolyl-isomerase and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 88 to 735 of SEQ ID NO: 9.

SEQ ID NO: 12

In a particular embodiment the polynucleotide of the invention encodes a acid phosphatase or a phytase or a phospholipase C and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 82 to 1824 of SEQ ID NO: 12.

SEQ ID NO: 13 and 14

In a particular embodiment the polynucleotide of the invention encodes a polysaccharide deacetylase or a xylan deacetylase and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 76 to 750 of SEQ ID NO: 13 or position 64 to 972 of SEQ ID NO: 14.

SEQ ID NO: 15

In a particular embodiment the polynucleotide of the invention encodes a sulfite oxidase and comprises or consists of an nucleotide sequence which has at least 70% identity, more particu-

larly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 88 to 642 of SEQ ID NO: 15.

SEQ ID NO: 16

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 64 to 771 of SEQ ID NO: 16.

SEQ ID NO: 17

In a particular embodiment the polynucleotide of the invention encodes mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 73 to 3390 of SEQ ID NO: 17.

SEQ ID NO: 18

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 124 to 744 of SEQ ID NO: 18.

SEQ ID NO: 19

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97%

identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 76 to 516 of SEQ ID NO: 19.

SEQ ID NO: 20

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 91 to 726 of SEQ ID NO: 20.

SEQ ID NO: 21

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 73 to 540 of SEQ ID NO: 21.

SEQ ID NO: 22

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 76 to 1431 of SEQ ID NO: 22.

SEQ ID NO: 23

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most

particularly 100% identity with the nucleotide sequence of position 58 to 1020 of SEQ ID NO: 23.

SEQ ID NO: 24

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 88 to 1023 of SEQ ID NO: 24.

SEQ ID NO: 25

In a particular embodiment the polynucleotide of the invention encodes a mature functional polypeptide and comprises or consists of an nucleotide sequence which has at least 70% identity, more particularly at least 80% identity, more particularly at least 90% identity, more particularly at least 95% identity, more particularly at least 96% identity, more particularly at least 97% identity, more particularly at least 98% identity, more particularly at least 99% identity or most particularly 100% identity with the nucleotide sequence of position 85 to 1197 of SEQ ID NO: 25.

Modification of a nucleotide sequence encoding a polypeptide of the present invention may be necessary for the synthesis of a polypeptide which comprises an amino acid sequence that has at least one substitution, deletion and/or insertion as compared to an amino acid sequence selected from mature polypeptide comprised in SEQ ID NO: 26 to SEQ ID NO: 50.

It will be apparent to those skilled in the art that such modifications can be made to preserve the function of the enzyme i.e. made outside regions critical to the function of the enzyme. Amino acid residues which are essential to the function are therefore preferably not subject to modification, such as substitution. Amino acid residues essential to the function may be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (see, *e.g.*, Cunningham and Wells, 1989, *Science* 244: 1081-1085). Sites of substrate-enzyme interaction can be determined by analysis of the three-dimensional structure as determined by such techniques as nuclear magnetic resonance analysis, crystallography or photoaffinity labeling (see, *e.g.*, de Vos *et al.*, 1992, *Science* 255: 306-312; Smith *et al.*,

1992, *Journal of Molecular Biology* 224: 899-904; Wlodaver *et al.*, 1992, *FEBS Letters* 309: 59-64).

Moreover, a nucleotide sequence encoding an enzyme of the invention may be modified by introduction of nucleotide substitutions which do not give rise to another amino acid sequence of the enzyme encoded by the nucleotide sequence, but which correspond to the codon usage of the host organism intended for production of the enzyme.

The introduction of a mutation into the nucleotide sequence to exchange one nucleotide for another nucleotide may be accomplished by site-directed mutagenesis using any of the methods known in the art. Particularly useful is the procedure, which utilizes a super coiled, double stranded DNA vector with an insert of interest and two synthetic primers containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, extend during temperature cycling by means of *Pfu* DNA polymerase. On incorporation of the primers, a mutated plasmid containing staggered nicks is generated. Following temperature cycling, the product is treated with *Dpn*I, which is specific for methylated and hemimethylated DNA to digest the parental DNA template and to select for mutation-containing synthesized DNA. Other procedures known in the art may also be used. For a general description of nucleotide substitution, one may consult with *e.g.*, Ford *et al.*, 1991, *Protein Expression and Purification* 2: 95-107.

The present invention also relates to a polynucleotide comprising, preferably consisting of, a nucleotide sequence which encodes a polypeptide of the invention and which hybridizes under high stringency conditions, preferably under very high stringency conditions with a polynucleotide probe selected from the group consisting of:

(i) the complementary strand to a nucleotide sequence selected from the regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide ,

(ii) the complementary strand to the cDNA sequence contained in a nucleotide sequences selected from the regions of SEQ ID NO: 1 to SEQ ID NO: 25 encoding a mature polypeptide and,

(iii) a fragment of (i) or (ii) encoding a secreted mature polypeptide having the function of the corresponding mature polypeptides comprised in SEQ ID NO:26 to SEQ ID NO:50

(J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York).

As will be understood, details and particulars concerning hybridization of the nucleotide sequences will be the same or analogous to the hybridization aspects discussed in the section titled "polypeptides of the invention" herein.

The present invention also encompasses a storage medium suitable for use in an electronic device comprising information of the amino acid sequence of polypeptides of the invention or the nucleotide sequences of the polynucleotide of the invention. The storage medium may suitably be a magnetic or optical disk and the electronic device a computing device and the information may in particular be stored on the storage medium in a digital form.

Nucleotide constructs

The present invention also relates to nucleic acid constructs comprising a nucleotide sequence of the invention operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

A nucleotide sequence encoding an enzyme of the invention may be manipulated in a variety of ways to provide for expression of the enzyme. Manipulation of the nucleotide sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying nucleotide sequences utilizing recombinant DNA methods are well known in the art.

The control sequence may be an appropriate promoter sequence, a nucleotide sequence that is recognized by a host cell for expression of the nucleotide sequence. The promoter sequence contains transcriptional control sequences, which mediate the expression of the polypeptide. The promoter may be any nucleotide sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extra cellular or intracellular polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, especially in a bacterial host cell, are the promoters obtained from the *E. coli lac* operon, *Streptomyces coelicolor* agarase gene (*dagA*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus subtilis xylA* and *xylB* genes, and prokaryotic beta-lactamase gene (Villa-Kamaroff *et al.*, 1978, *Proceedings of the National Academy of Sciences USA* 75: 3727-3731), as well as the *tac* promoter (DeBoer

et al., 1983, *Proceedings of the National Academy of Sciences USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242: 74-94; and in Sambrook *et al.*, 1989, *supra*.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, and *Fusarium oxysporum* trypsin-like protease (WO 96/00787), as well as the NA2-tpi promoter (a hybrid of the promoters from the genes for *Aspergillus niger* neutral alpha-amylase and *Aspergillus oryzae* triose phosphate isomerase), and mutant, truncated, and hybrid promoters thereof.

In a yeast host, useful promoters are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos *et al.*, 1992, *Yeast* 8: 423-488.

The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3' terminus of the nucleotide sequence encoding the enzyme. Any terminator which is functional in the host cell of choice may be used in the present invention.

Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* alpha-glucosidase, and *Fusarium oxysporum* trypsin-like protease.

Preferred terminators for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), and *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos *et al.*, 1992, *supra*.

The control sequence may also be a suitable leader sequence, a non-translated region of an mRNA which is important for translation by the host cell. The leader sequence is operably linked to the 5' terminus of the nucleotide sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used in the present invention.

Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3' terminus of the nucleotide sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention.

Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Fusarium oxysporum* trypsin-like protease, and *Aspergillus niger* alpha-glucosidase.

Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, *Molecular Cellular Biology* 15: 5983-5990.

The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded enzyme into the cell's secretory pathway. The 5' end of the coding sequence of the nucleotide sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted enzyme. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region which is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to enhance secretion of the enzyme. However, any signal peptide coding region which directs the expressed enzyme into the secretory pathway of a host cell of choice may be used in the present invention.

Effective signal peptide coding regions for bacterial host cells are the signal peptide coding regions obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* neutral proteases (*nprT*, *nprS*, *nprM*), and *Bacillus sub-*

tilis prsA. Further signal peptides are described by Simonen and Palva, 1993, *Microbiological Reviews* 57: 109-137.

Effective signal peptide coding regions for filamentous fungal host cells are the signal peptide coding regions obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Rhizomucor miehei* aspartic proteinase, *Humicola insolens* cellulase, and *Humicola lanuginosa* lipase.

Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding regions are described by Romanos *et al.*, 1992, *supra*.

The control sequence may also be a propeptide coding region that codes for an amino acid sequence positioned at the amino terminus of a enzyme. The resultant polypeptide may be denoted a pro-enzyme or propolypeptide. A propolypeptide is generally inactive and can be converted to a mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the genes for *Bacillus subtilis* alkaline protease (*aprE*), *Bacillus subtilis* neutral protease (*nprT*), *Saccharomyces cerevisiae* alpha-factor, *Rhizomucor miehei* aspartic proteinase, and *Myceliophthora thermophila* laccase (WO 95/33836.

Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.
In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the TAKA alpha-amylase promoter, *Aspergillus niger* glucoamylase promoter, and *Aspergillus oryzae* glucoamylase promoter may be used as regulatory sequences.

Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene which is amplified in the presence of methotrexate, and the metallothionein genes which are amplified with heavy metals. In these cases, the nucleotide sequence encoding the polypeptide would be operably linked with the regulatory sequence.

Recombinant expression vectors

The present invention also relates to recombinant expression vectors comprising the nucleic acid construct of the invention. The various nucleotide and control sequences described above may be joined together to produce a recombinant expression vector, which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleotide

sequence encoding the polypeptide at such sites. Alternatively, the nucleotide sequence of the present invention may be expressed by inserting the nucleotide sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (*e.g.*, a plasmid or virus) that can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleotide sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

The vector may be an autonomously replicating vector, *i.e.*, a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome.

The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

The vectors of the present invention preferably contain one or more selectable markers that permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

Examples of bacterial selectable markers are the *dal* genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers that confer antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracycline resistance. Suitable markers for yeast host cells are ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. Selectable markers for use in a filamentous fungal host cell include, but are not limited to, *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hygB* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenyltransferase), *trpC* (anthranilate synthase), as well as equivalents thereof.
Preferred for use in an *Aspergillus* cell are the *amdS* and *pyrG* genes of *Aspergillus nidulans* or *Aspergillus oryzae* and the *bar* gene of *Streptomyces hygroscopicus*.

The vectors of the present invention preferably contain an element(s) that permits stable integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

For integration into the host cell genome, the vector may rely on the nucleotide sequence encoding the polypeptide or any other element of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleotide sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleotide sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleotides, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to 1,500 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleotide sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAMß1 permitting replication in *Bacillus*. Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6.

The origin of replication may be one having a mutation which makes its functioning temperature-sensitive in the host cell (see, *e.g.*, Ehrlich, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1433).

More than one copy of a nucleotide sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the nucleotide sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleotide sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleotide sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

Recombinant host cells

The present invention also relates to recombinant a host cell comprising the nucleic acid construct of the invention, which are advantageously used in the recombinant production of the polypeptides. A vector comprising a nucleotide sequence of the present invention is introduced into a host cell so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier.

The host cell may be a unicellular microorganism, *e.g.*, a prokaryote or a non-unicellular microorganism, *e.g.*, a eukaryote.

Useful unicellular cells are bacterial cells such as gram positive bacteria including, but not limited to, a *Bacillus* cell, *e.g.*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis*; or a *Streptomyces* cell, *e.g.*, *Streptomyces lividans* or *Streptomyces murinus*, or gram negative bacteria such as *E. coli* and *Pseudomonas* sp. In a preferred embodiment, the bacterial host cell is a *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus*, or *Bacillus subtilis* cell. In another preferred embodiment, the *Bacillus* cell is an alkalophilic *Bacillus*.

The introduction of a vector into a bacterial host cell may, for instance, be effected by protoplast transformation (see, *e.g.*, Chang and Cohen, 1979, *Molecular General Genetics* 168: 111-115), using competent cells (see, *e.g.*, Young and Spizizin, 1961, *Journal of Bacteriology* 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209-221), electroporation (see, *e.g.*, Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), or conjugation (see, *e.g.*, Koehler and Thorne, 1987, *Journal of Bacteriology* 169: 5771-5278).

The host cell may be a eukaryote, such as a mammalian, insect, plant, or fungal cell.

In a preferred embodiment, the host cell is a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (as defined by Hawksworth *et al.*, *In*, *Ainsworth and Bisby's Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the Oomycota (as cited in Hawksworth *et al.*, 1995, *supra*, page 171) and all mitosporic fungi (Hawksworth *et al.*, 1995, *supra*). In a more preferred embodiment, the fungal host cell is a yeast cell. "Yeast" as used herein includes

ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast belonging to the Fungi Imperfecti (Blastomycetes). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in *Biology and Activities of Yeast* (Skinner, F.A., Passmore, S.M., and Davenport, R.R., eds, *Soc. App. Bacteriol. Symposium Series* No. 9, 1980).

In an even more preferred embodiment, the yeast host cell is a *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cell.

In a most preferred embodiment, the yeast host cell is a *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis* or *Saccharomyces oviformis* cell. In another most preferred embodiment, the yeast host cell is a *Kluyveromyces lactis* cell. In another most preferred embodiment, the yeast host cell is a *Yarrowia lipolytica* cell.

In another more preferred embodiment, the fungal host cell is a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.*, 1995, *supra*). The filamentous fungi are characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

In an even more preferred embodiment, the filamentous fungal host cell is a cell of a species of, but not limited to, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium*, or *Trichoderma*.

In a most preferred embodiment, the filamentous fungal host cell is an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* or *Aspergillus oryzae* cell. In another most preferred embodiment, the filamentous fungal host cell is a *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, or *Fusarium venenatum* cell. In an even most preferred embodiment, the filamentous fungal parent cell is a *Fusarium venenatum* (Nirenberg sp. nov.) cell. In another most preferred embodiment, the filamentous fungal host cell is a *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicil-*

lium purpurogenum, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* host cells are described in EP 238 023 and Yelton *et al.*, 1984, *Proceedings of the National Academy of Sciences USA* 81: 1470-1474. Suitable methods for transforming *Fusarium* species are described by Malardier *et al.*, 1989, *Gene* 78: 147-156 and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, *In* Abelson, J.N. and Simon, M.I., editors, *Guide to Yeast Genetics and Molecular Biology*, *Methods in Enzymology*, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito *et al.*, 1983, *Journal of Bacteriology* 153: 163; and Hinnen *et al.*, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1920.

Methods for preparing enzyme polypeptides

The present invention also relates to methods for producing an enzyme of the invention comprising (a) cultivating a strain comprising a nucleotide sequence encoding an enzyme of the invention which strain is capable of expressing and secreting the enzyme and (b) recovering the enzyme. In a particular embodiment the strain is a wild type strain such as the *Alicyclobacillus sp.* DSM 15716, while in another embodiment the strain is a recombinant host cell as described, *supra*.

In these methods of the invention, the cells are cultivated in a nutrient medium suitable for production of the enzyme using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (*e.g.*, in catalogues of the American Type Culture Collection). As the enzyme is secreted into the nutrient medium, the enzyme can be recovered directly from the medium.

The resulting enzyme may be recovered by methods known in the art. For example, the enzyme may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

The polypeptides of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (*e.g.*, ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.*, preparative isoelectric focusing), differential solubility (*e.g.*, ammonium sulfate precipitation), SDS-PAGE, or extraction (see, *e.g.*, *Protein Purification*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

Transgenic plants

The present invention also relates to a transgenic plant, plant part, or plant cell that has been transformed with a nucleotide sequence encoding an enzyme of the invention so as to express and produce the enzyme. In one embodiment the plant could be used as host for production of enzyme in recoverable quantities. The enzyme may be recovered from the plant or plant part. Alternatively, the plant or plant part containing the recombinant enzyme may be used as such for improving the quality of a food or feed, *e.g.*, improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor. In particular the plant or plant parts expressing the enzyme may be used as an improved starting material for production of fuel-alcohols or bio-ethanol

The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot). Examples of monocot plants are grasses, such as meadow grass (blue grass, Poa), forage grass such as festuca, lolium, temperate grass, such as Agrostis, and cereals, *e.g.*, wheat, oats, rye, barley, rice, sorghum, and maize (corn).

Examples of dicot plants are tobacco, legumes, such as lupins, potato, sugar beet, pea, bean and soybean, and cruciferous plants (family Brassicaceae), such as cauliflower, rape seed, and the closely related model organism *Arabidopsis thaliana*.

Examples of plant parts are stem, callus, leaves, root, fruits, seeds, and tubers. Also specific plant tissues, such as chloroplast, apoplast, mitochondria, vacuole, peroxisomes, and cytoplasm are considered to be a plant part. Furthermore, any plant cell, whatever the tissue origin, is considered to be a plant part.

Also included within the scope of the present invention are the progeny of such plants, plant parts and plant cells.

The transgenic plant or plant cell expressing an enzyme of the invention may be constructed in accordance with methods known in the art. Briefly, the plant or plant cell is constructed by incorporating one or more expression constructs encoding an enzyme of the

invention into the plant host genome and propagating the resulting modified plant or plant cell into a transgenic plant or plant cell.

Conveniently, the expression construct is a nucleic acid construct which comprises a nucleotide sequence encoding an enzyme of the present invention operably linked with appropriate regulatory sequences required for expression of the nucleotide sequence in the plant or plant part of choice. Furthermore, the expression construct may comprise a selectable marker useful for identifying host cells into which the expression construct has been integrated and DNA sequences necessary for introduction of the construct into the plant in question (the latter depends on the DNA introduction method to be used).

The choice of regulatory sequences, such as promoter and terminator sequences and optionally signal or transit sequences, is determined, for example, on the basis of when, where, and how the enzyme is desired to be expressed. For instance, the expression of the gene encoding an enzyme of the invention may be constitutive or inducible, or may be developmental, stage or tissue specific, and the gene product may be targeted to a specific tissue or plant part such as seeds or leaves. Regulatory sequences are, for example, described by Tague *et al.*, 1988, *Plant Physiology* 86: 506.

For constitutive expression, the 35S-CaMV promoter may be used (Franck *et al.*, 1980, *Cell* 21: 285-294). Organ-specific promoters may be, for example, a promoter from storage sink tissues such as seeds, potato tubers, and fruits (Edwards & Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), or from metabolic sink tissues such as meristems (Ito *et al.*, 1994, *Plant Mol. Biol.* 24: 863-878), a seed specific promoter such as the glutelin, prolamin, globulin, or albumin promoter from rice (Wu *et al.*, 1998, *Plant and Cell Physiology* 39: 885-889), a *Vicia faba* promoter from the legumin B4 and the unknown seed protein gene from *Vicia faba* (Conrad *et al.*, 1998, *Journal of Plant Physiology* 152: 708-711), a promoter from a seed oil body protein (Chen *et al.*, 1998, Plant and Cell Physiology 39: 935-941), the storage protein napA promoter from *Brassica napus*, or any other seed specific promoter known in the art, *e.g.*, as described in WO 91/14772. Furthermore, the promoter may be a leaf specific promoter such as the *rbcs* promoter from rice or tomato (Kyozuka *et al.*, 1993, *Plant Physiology* 102: 991-1000, the chlorella virus adenine methyltransferase gene promoter (Mitra and Higgins, 1994, *Plant Molecular Biology* 26: 85-93), or the *aldP* gene promoter from rice (Kagaya *et al.*, 1995, *Molecular and General Genetics* 248: 668-674), or a wound inducible promoter such as the potato pin2 promoter (Xu *et al.*, 1993, *Plant Molecular Biology* 22: 573-588).

A promoter enhancer element may also be used to achieve higher expression of the enzyme of the invention in the plant. For instance, the promoter enhancer element may be an

intron which is placed between the promoter and the nucleotide sequence encoding an enzyme of the present invention. For instance, Xu *et al.*, 1993, *supra* disclose the use of the first intron of the rice actin 1 gene to enhance expression.

The selectable marker gene and any other parts of the expression construct may be chosen from those available in the art.

The nucleic acid construct is incorporated into the plant genome according to conventional techniques known in the art, including *Agrobacterium*-mediated transformation, virus-mediated transformation, microinjection, particle bombardment, biolistic transformation, and electroporation (Gasser *et al.*, 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto *et al.*, 1989, *Nature* 338: 274).

Presently, *Agrobacterium tumefaciens*-mediated gene transfer is the method of choice for generating transgenic dicots (for a review, see Hooykas and Schilperoort, 1992, *Plant Molecular Biology* 19: 15-38). However it can also be used for transforming monocots, although other transformation methods are generally preferred for these plants. Presently, the method of choice for generating transgenic monocots is particle bombardment (microscopic gold or tungsten particles coated with the transforming DNA) of embryonic calli or developing embryos (Christou, 1992, *Plant Journal* 2: 275-281; Shimamoto, 1994, *Current Opinion Biotechnology* 5: 158-162; Vasil *et al.*, 1992, *Bio/Technology* 10: 667-674). An alternative method for transformation of monocots is based on protoplast transformation as described by Omirulleh *et al.*, 1993, *Plant Molecular Biology* 21: 415-428.

Following transformation, the transformants having incorporated therein the expression construct are selected and regenerated into whole plants according to methods well known in the art.

The present invention also relates to methods for producing an enzyme of the invention comprising (a) cultivating a transgenic plant or a plant cell comprising a nucleotide sequence encoding an enzyme of the invention under conditions conducive for production of the enzyme and (b) recovering the enzyme.

Compositions comprising polypeptides and methods for their preparation

The invention provide a composition comprising a polypeptide of the invention and an excipient and a method for preparing such a composition comprising admixing the polypeptide of the invention with an excipient. In a particular embodiment the polypeptide of the invention is the major (polypeptide) component of the composition, *e.g.*, a mono-component composition. The ex-

cipient in this context is to be understood as any auxilliary agent or compound used to formulate the composition and includes solvent, carriers, stabilizers and the like.

The composition may further comprise one or more additional enzymes, such as an aminopeptidase, amylase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, alpha-galactosidase, beta-galactosidase, glucoamylase, alpha-glucosidase, beta-glucosidase, haloperoxidase, invertase, laccase, lipase, mannosidase, oxidase, pectinolytic enzyme, peptidoglutaminase, peroxidase, phytase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, or xylanase.

The compositions may be prepared in accordance with methods known in the art and may be in the form of a liquid or a solid composition. For instance, the enzyme composition may be formulated using methods known to the art of formulating polypeptides and/or pharmaceutical products, e.g. into coated or uncoated granules or micro-granules. The polypeptide of the invention may thus be provided in the form of a granule, preferably a non-dusting granule, a liquid, in particular a stabilized liquid, a slurry or a protected polypeptide. For certain applications, immobilization of the polypeptide on a solid matrix may be preferred.

The polypeptide to be included in the composition may be stabilized in accordance with methods known in the art e.g. by stabilizing the polypeptide in the composition by adding and antioxidant or reducing agent to limit oxidation of the polypeptide or it may be stabilized by adding polymers such as PVP, PVA, PEG or other suitable polymers known to be beneficial to the stability of polypeptides in solid or liquid compositions

In a further embodiment the composition of the invention is a detergent composition which, in addition to the polypeptide of the invention, comprises a surfactant and optionally compounds selected from the group consisting of builders such as zeolites, bleaching agents such as percarbonate, bleach enhancers such as TAED or NOBS, suds suppressors, fragrants, etc.

In a further embodiment the composition of the invention is a feed composition that in addition to the polypeptide of the invention comprises a cereal or grain product.

In a further embodiment the composition of the invention is a food composition such as a bakers flour composition, a brewed product, a fruit juice, an oil or lard product comprising the polypeptide of the invention.

In a further embodiment the composition of the invention is a pulping composition, which in addition to the polypeptide of the invention, comprises pulp.

In a further embodiment the composition of the invention is a biocidal composition, which comprises in addition to the polypeptide of the invention, an oxidoreductase enhancer.

Use of polypeptides or compositions comprising them

In still further aspects the invention provides use of the polypeptides or polynucleotides of the invention or a composition comprising said polypeptides or polynucleotides in various applications, particularly (technical) processes such as processes performed in industry or household, herein under for commercial research purposes. Hence the invention encompasses a process comprising employing a polypeptide of the invention or a polynucleotide of the invention in a (technical) industrial, research or household process.

In one embodiment the polypeptide or the composition of the invention is used for cleaning a cellulosic fabric.

In another embodiment the polypeptide or the composition of the invention is used to prepare a food or feed additive.

In yet another embodiment the polypeptide or the composition of the invention is used for treatment of lignolosic materials and pulp.

DETERGENT DISCLOSURE

The polypeptide of the invention may be added to and thus become a component of a detergent composition.

The detergent composition of the invention may for example be formulated as a hand or machine laundry detergent composition including a laundry additive composition suitable for pre-treatment of stained fabrics and a rinse added fabric softener composition, or be formulated as a detergent composition for use in general household hard surface cleaning operations, or be formulated for hand or machine dishwashing operations.

In a specific aspect, the invention provides a detergent additive comprising the polypeptide of the invention. The detergent additive as well as the detergent composition may comprise one or more other enzymes such as a protease, a lipase, a cutinase, an amylase, a carbohydrase, a cellulase, a pectinase, a mannanase, an arabinase, a galactanase, a xylanase, an oxidase, e.g., a laccase, and/or a peroxidase.

In general the properties of the chosen enzyme(s) should be compatible with the selected detergent, (i.e. pH-optimum, compatibility with other enzymatic and non-enzymatic ingredients, etc.), and the enzyme(s) should be present in effective amounts.

Proteases: Suitable proteases include those of animal, vegetable or microbial origin. Microbial origin is preferred. Chemically modified or protein engineered mutants are included. The protease may be a serine protease or a metallo protease, preferably an alkaline microbial

protease or a trypsin-like protease. Examples of alkaline proteases are subtilisins, especially those derived from Bacillus, e.g., subtilisin Novo, subtilisin Carlsberg, subtilisin 309, subtilisin 147 and subtilisin 168 (described in WO 89/06279). Examples of trypsin-like proteases are trypsin (e.g. of porcine or bovine origin) and the Fusarium protease described in WO 89/06270 and WO 94/25583.

Examples of useful proteases are the variants described in WO 92/19729, WO 98/20115, WO 98/20116, and WO 98/34946, especially the variants with substitutions in one or more of the following positions: 27, 36, 57, 76, 87, 97, 101, 104, 120, 123, 167, 170, 194, 206, 218, 222, 224, 235 and 274.

Preferred commercially available protease enzymes include Alcalase®, Savinase®, Primase®, Duralase®, Esperase®, and Kannase® (Novozymes A/S), Maxatase®, Maxacal®, Maxapem®, Properase®, Purafect®, Purafect OxP®, FN2®, and FN3® (Genencor International Inc.).

Lipases: Suitable lipases include those of bacterial or fungal origin. Chemically modified or protein engineered mutants are included. Examples of useful lipases include lipases from Humicola (synonym Thermomyces), e.g. from H. lanuginosa (T. lanuginosus) as described in EP 258 068 and EP 305 216 or from H. insolens as described in WO 96/13580, a Pseudomonas lipase, e.g. from P. alcaligenes or P. pseudoalcaligenes (EP 218 272), P. cepacia (EP 331 376), P. stutzeri (GB 1,372,034), P. fluorescens, Pseudomonas sp. strain SD 705 (WO 95/06720 and WO 96/27002), P. wisconsinensis (WO 96/12012), a Bacillus lipase, e.g. from B. subtilis (Dartois et al. (1993), Biochemica et Biophysica Acta, 1131, 253-360), B. stearothermophilus (JP 64/744992) or B. pumilus (WO 91/16422).

Other examples are lipase variants such as those described in WO 92/05249, WO 94/01541, EP 407 225, EP 260 105, WO 95/35381, WO 96/00292, WO 95/30744, WO 94/25578, WO 95/14783, WO 95/22615, WO 97/04079 and WO 97/07202.

Preferred commercially available lipase enzymes include LipolaseTM, Lipolase UltraTM and Lipex (Novozymes A/S).

Amylases: Suitable amylases (alpha and/or beta) include those of bacterial or fungal origin. Chemically modified or protein engineered mutants are included. Amylases include, for example, alpha-amylases obtained from Bacillus, e.g. a special strain of B. licheniformis, described in more detail in GB 1,296,839.

Examples of useful amylases are the variants described in WO 94/02597, WO 94/18314, WO 96/23873, and WO 97/43424, especially the variants with substitutions in one or

more of the following positions: 15, 23, 105, 106, 124, 128, 133, 154, 156, 181, 188, 190, 197, 202, 208, 209, 243, 264, 304, 305, 391, 408, and 444.

Commercially available amylases are Duramyl™, Termamyl™, Fungamyl™ and BAN™ (Novozymes A/S), Rapidase™ and Purastar™ (from Genencor International Inc.).

Cellulases: Suitable cellulases include those of bacterial or fungal origin. Chemically modified or protein engineered mutants are included. Suitable cellulases include cellulases from the genera Bacillus, Pseudomonas, Humicola, Fusarium, Thielavia, Acremonium, e.g. the fungal cellulases produced from Humicola insolens, Myceliophthora thermophila and Fusarium oxysporum disclosed in US 4,435,307, US 5,648,263, US 5,691,178, US 5,776,757 and WO 89/09259.

Especially suitable cellulases are the alkaline or neutral cellulases having colour care benefits. Examples of such cellulases are cellulases described in EP 0 495 257, EP 0 531 372, WO 96/11262, WO 96/29397, WO 98/08940. Other examples are cellulase variants such as those described in WO 94/07998, EP 0 531 315, US 5,457,046, US 5,686,593, US 5,763,254, WO 95/24471, WO 98/12307 and PCT/DK98/00299.

Commercially available cellulases include Celluzyme®, and Carezyme® (Novozymes), Clazinase®, and Puradax HA® (Genencor International Inc.), and KAC-500(B) ® (Kao Corporation).

Peroxidases/Oxidases: Suitable peroxidases/oxidases include those of plant, bacterial or fungal origin. Chemically modified or protein engineered mutants are included. Examples of useful peroxidases include peroxidases from Coprinus, e.g. from C. cinereus, and variants thereof as those described in WO 93/24618, WO 95/10602, and WO 98/15257.

Commercially available peroxidases include Guardzyme® (Novozymes A/S).

The detergent enzyme(s) may be included in a detergent composition by adding separate additives containing one or more enzymes, or by adding a combined additive comprising all of these enzymes. A detergent additive of the invention, i.e. a separate additive or a combined additive, can be formulated e.g. as a granulate, a liquid, a slurry, etc. Preferred detergent additive formulations are granulates, in particular non-dusting granulates, liquids, in particular stabilized liquids, or slurries.

Non-dusting granulates may be produced, e.g., as disclosed in US 4,106,991 and 4,661,452 and may optionally be coated by methods known in the art. Examples of waxy coating materials are poly(ethylene oxide) products (polyethyleneglycol, PEG) with mean molar weights of 1000 to 20000; ethoxylated nonylphenols having from 16 to 50 ethylene oxide units; ethoxylated fatty alcohols in which the alcohol contains from 12 to 20 carbon atoms and in

which there are 15 to 80 ethylene oxide units; fatty alcohols; fatty acids; and mono- and di- and triglycerides of fatty acids. Examples of film-forming coating materials suitable for application by fluid bed techniques are given in GB 1483591. Liquid enzyme preparations may, for instance, be stabilized by adding a polyol such as propylene glycol, a sugar or sugar alcohol, lactic acid or boric acid according to established methods. Protected enzymes may be prepared according to the method disclosed in EP 238,216.

The detergent composition of the invention may be in any convenient form, e.g., a bar, a tablet, a powder, a granule, a paste or a liquid. A liquid detergent may be aqueous, typically containing up to 70 % water and 0-30 % organic solvent, or non-aqueous.

The detergent composition comprises one or more surfactants, which may be non-ionic including semi-polar and/or anionic and/or cationic and/or zwitterionic. The surfactants are typically present at a level of from 0.1% to 60% by weight.

When included therein the detergent will usually contain from about 1% to about 40% of an anionic surfactant such as linear alkylbenzenesulfonate, alpha-olefinsulfonate, alkyl sulfate (fatty alcohol sulfate), alcohol ethoxysulfate, secondary alkanesulfonate, alpha-sulfo fatty acid methyl ester, alkyl- or alkenylsuccinic acid or soap.

When included therein the detergent will usually contain from about 0.2% to about 40% of a non-ionic surfactant such as alcohol ethoxylate, nonylphenol ethoxylate, alkyl-polyglycoside, alkyldimethylamineoxide, ethoxylated fatty acid monoethanolamide, fatty acid monoethanolamide, polyhydroxy alkyl fatty acid amide, or N-acyl N-alkyl derivatives of glucosamine ("glucamides").

The detergent may contain 0-65 % of a detergent builder or complexing agent such as zeolite, diphosphate, triphosphate, phosphonate, carbonate, citrate, nitrilotriacetic acid, ethylenediaminetetraacetic acid, diethylenetriaminepentaacetic acid, alkyl- or alkenylsuccinic acid, soluble silicates or layered silicates (e.g. SKS-6 from Hoechst).

The detergent may comprise one or more polymers. Examples are carboxy-methylcellulose, poly(vinylpyrrolidone), poly (ethylene glycol), poly(vinyl alcohol), poly(vinylpyridine-N-oxide), poly(vinylimidazole), polycarboxylates such as polyacrylates, maleic/acrylic acid copolymers and lauryl methacrylate/acrylic acid copolymers.

The detergent may contain a bleaching system which may comprise a H2O2 source such as perborate or percarbonate which may be combined with a peracid-forming bleach activator such as tetraacetylethylenediamine or nonanoyloxybenzenesulfonate. Alternatively, the bleaching system may comprise peroxyacids of e.g. the amide, imide, or sulfone type.

The enzyme(s) of the detergent composition of the invention may be stabilized using conventional stabilizing agents, e.g., a polyol such as propylene glycol or glycerol, a sugar or sugar alcohol, lactic acid, boric acid, or a boric acid derivative, e.g., an aromatic borate ester, or a phenyl boronic acid derivative such as 4-formylphenyl boronic acid, and the composition may be formulated as described in e.g. WO 92/19709 and WO 92/19708.

The detergent may also contain other conventional detergent ingredients such as e.g. fabric conditioners including clays, foam boosters, suds suppressors, anti-corrosion agents, soil-suspending agents, anti-soil redeposition agents, dyes, bactericides, optical brighteners, hydrotropes, tarnish inhibitors, or perfumes.

It is at present contemplated that in the detergent compositions any enzyme, in particular the enzyme of the invention, may be added in an amount corresponding to 0.01-100 mg of enzyme protein per litre of wash liquor, preferably 0.05-5 mg of enzyme protein per liter of wash liquor, in particular 0.1-1 mg of enzyme protein per litre of wash liquor.

The enzyme of the invention may additionally be incorporated in the detergent formulations disclosed in WO 97/07202 that is hereby incorporated as reference.

DEPOSITED MICROORGANISMS

The following microorganism were deposited by the applicant according to the Budapest Treaty on the International Recognition of the Deposits of Microorganisms for the Purpose of Patent Procedures at Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig, Germany:

June 30, 2003: *Alicyclobacillus sp.* CS81 thermo-acidophile; DSM accession No. 15716

EXAMPLES

Example 1 Identifying functional polypeptides secreted by *Alicyclobacillus sp.* DSM 15716

Genomic library construction

Chromosomal DNA from *Alicyclobacillus sp.* DSM 15716 was prepared by using standard molecular biology techniques (Ausuble et al. 1995 "Current protocols in molecular biology" Publ: John Wiley and sons). The prepared DNA was partially cleaved with Sau3A and separated on

an agarose gel. Fragments of 3 to 8 kilobases were eluted and precipitated and resuspended in a suitable buffer.

A genomic library was made by using the Stratagene ZAP Express ™ predigested Vector kit and Stratagene ZAP Express ™predigested Gigapack ® cloning kit (Bam HI predigested) (Stratagene Inc., USA) following the instructions/recommendations from the vendor. The resulting lambdaZAP library comprised 38000 pfu of which 10000 were collected for mass excision. The resulting 70000 E. coli colonies were pooled and plasmids were prepared by using the Qiagen Spin Mini prep kit (Qiagen, Germany). The eluate of approx. 1 ml containing the plasmid DNA was precipitated in a centrifuge with 1 volume part of Na-acetate pH5 and 2 volume parts 96% ethanol at 20000 rpm at 4C, washed with 70% v/v ethanol, dried at room temperature and resuspended in 200 microl TE buffer. The DNA concentration of the plasmid pool DNA of the Alicyclobacillus sp. genomic library was 5.2 microgram/microliter.

Transposon construction and preparation

The rationale behind the methology of Transposon Assisted Signal Trapping (TAST) as described in WO 01/77315 A1 is to fuse all genes within a selected genome with a gene encoding a signalless beta-lactamase via a transposon tag. Hence when growing host cell clones comprising the genes of a genome fused with a gene encoding a signalless beta-lactamase via a transposon tag in an ampicillin containg medium only those clones expressing and secreting a beta-lactamase will survive. However the beta-lactamase will only be secreted if the gene to which the beta-lactamase gene is fused has an intact promotor and ribosome binding site (i.e. a gene which is expressed by the cell to produce a polypeptide in real life), which can be recognized in the host strain, and if the beta-lactamase is translated so that the synthesized polypeptide is transported across the cytoplasma membrane and folded correctly. Hence, when inserting the fused gene into a selected host cell, those clones, which are ampicillin resistant contains a gene which encodes a functional secreted polypeptide.

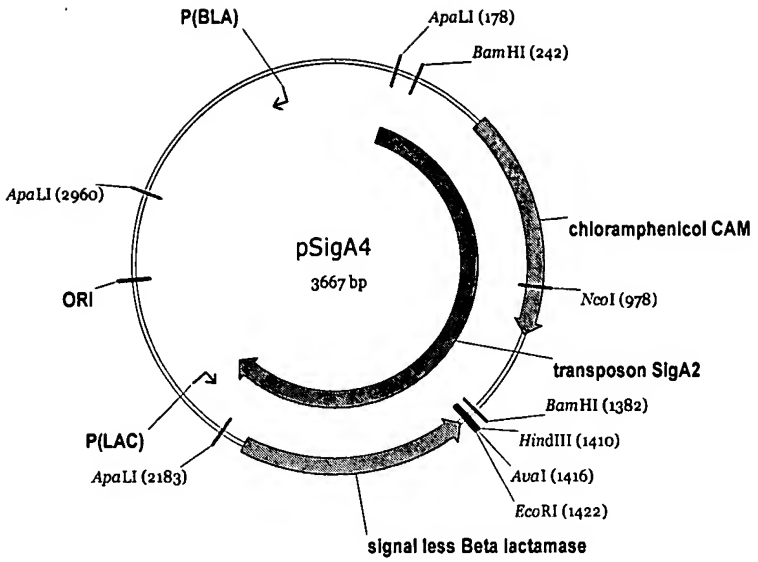
Usually, when employing the TAST methodology it is even not necessary to express the entire gene. When tagging the genes with a transposon, expression of the N-terminal part of the genes as protein fusion shows that the genes contain intact transcription, translation and secretion sequences. Hence expression of the N-terminal part of the genes as protein fusion is usually regarded as sufficient for assuring expression and secretion of the entire gene.

Thus it can be concluded that the genes obtained by the TAST method actually do encode secreted functional polypeptides.

Construction of a SigA4 transposon containing the β-lactamase reporter gene:

Following the instructions of WO 01/77315 A1, the construction of a transposon containing a signal-less β-lactamase gene was carried out using standard molecular biology techniques. The signal-less β-lactamase gene was initially PCR amplified from the vector pUC19) using a proof-reading polymerase (Pfu Turbo, Stratagene, USA). The resulting PCR fragment contained the restriction sites *Not*I and *Eco*RI in order to aid cloning. The plasmid pEntranceposon(Cam^r) containing the Entranceposon and the antibiotic resistance markers *CAT* (encoding chloramphencol resistance in the transposon) was obtained from Finnzymes, OY (Espoo Finland). The plasmid was digested with the restriction enzymes *Not*I and *Eco*RI, gel purified and ligated with the signal-less β-lactamase containing fragment. The ligation was transformed into electro-competent DH10B cells and the *E.coli* clone containing the recombinant plasmid with the signal-less β-lactamase was identified by restriction analysis and named SigA2.

For transposon preparation, a smaller derivative of SigA2 was constructed, which lacked the *bla* gene encoding beta-lactamase: Two oligonucleotide primers SigA2NotU-P 5'-TCG CGA TCC GTT TTC GCA TTT ATC GTG AAA CGC T-3' (SEQ ID NO: 51) and SigA2NotD-P 5'-CCG CAA ACG CTG GTG AAA GTA AAA GAT GCT GAA-3' (SEQ ID NO: 52), which bind to the start and stop of the *bla* gene of SigA2 directing outwards were used PCR amplify SigA2 without the *bla* gene. An amplificate of approx. 3,6 kb generated in the this PCR reaction was relegated and transformed in to a suitable E.coli strain. A plasmid of 3,6 kb was isolated from a transformant which was able to grow on LB chloramphenicol but not on LB ampicillin. This plasmid maintained both BglII sites and lacks the active *bla* gene and was called pSig4.



60 microliter of pSigA4 plasmid DNA preparation with a concentration of 0.3 microgram/microliter was digested with BglII and separated on an agarose gel. The SigA2 transposon DNA band of 2 kb was eluted and purified by using the "GFX™PCR, DNA and Gel Band Purification Kit" (Amersham Pharmacia Biotech Inc,USA) according to the instructions of the vender and eluted in 200 microliter EB buffer.

C. Transposon tagging

The transposon prepared from pSigA4 carries a 5'-truncated bla-gene encoding a β-lactamase from which the secretion signal has been removed. The β-lactamase conveys ampicillin resistance on *E. coli* only when the protein is secreted to the periplasm, whereas cytoplasmic expression of β-lactamase does not confer ampicillin resistance. Without a signal sequence, the β-lactamase enzyme will not be transported to the periplasm and therefore the clone will not grow on media containing ampicillin. The signal-less β-lactamase gene was contained within the transposon in such a way that there was a continuous open reading frame between the transposon border and the β-lactamase coding region. In this way the modified transposon, when it transposes into a gene encoding a protein that is secreted, could cause an in-frame fusion with the target gene. This resulted in a fusion gene product that is secreted to the periplasm of *E. coli* and conveys resistance to the ampicillin. If the transposon integrated even in-frame into a gene encoding a non-secreted protein, the respective host will not become ampicillin resistance.

For the in vitro transposon tagging of the Alicyclobacillus sp. library, 4 or 8 microliter of SigA2 transposon containing approx. 2,6 ug DNA were mixed with 1 microliter of the DNA con-

centration of the plasmid pool DNA of the Alicyclobacillus sp. genomic library, 2 microliter of Finnzymes MuA Transposase (0,22 microgram/microliter) and 5 microliter of 5x buffer from Finnzymes OY, Espoo, Finland) in a total volume of 50 microliter and incubated at 30 °C for 3,5 h and followed by heat inactivation at 75 °C for 10 min. The DNA was precipitated by addition of 5microliter 3M Na-acetate pH5 and 110 microliter 96% ethanol and centrifugation for 30 min at 20000 rpm. The pellet was washed and dried and resuspended in 10 microliter TE buffer.

D. Transformation and selection

Electro-competent *E. coli* DH10B cells were transformed by electroporation in a Biorad Gene Pulse device (50uF, 25mAmp, 1.8 kV with 5 microliter of the transposon tagged plasmid pool, mixed with 1ml SOC medium, pre-incubated for 1h at 37C and plated on LB with 25 microliter/mililiter ampicillin, 50 microliter/mililiter kanamycin, 10 microliter/mililiter chloramphenicol and incubated for 2-3 days. Out of the resistant transformants 1056 colonies were selected and plasmids were prepared by applying the Qiaprep 96 Turbo Biorobot kit according to the instructions of the vender.

E. Plasmid preparation and sequencing

1056 transposon tagged plasmids were sequenced in with the A2up primer AGCGTTTGCGGCCGCGATCC (SEQ ID NO: 53) which read upstream into the into the transposon tagged gene, and, in a second reaction, with B primer TTATTCGGTCGAAAAGGATCC (SEQ ID NO: 54) which read downstream into the transposon tagged gene.

F. Sequence assembly and annotation

The obtained sequences were assembled into contigs by using the program PhredPhrap (Brent Ewing, LaDeana Hillier, Michael C. Wendl, and Phil Green, Base-calling of automated sequencer traces using phred I. Accuracy assessment (1998) Genome Research 8:175-185; Brent Ewing and Phil Green, Base-calling of automated sequencer traces using phred II. Error probabilities (1998) Genome Research 8:186-194). The obtained contigs were subsequently compared to sequences available in standard public DNA and protein sequences databases by using the program BLASTX 2.0a19MP-WashU [14-Jul-1998] [Build linux-x86 18:51:44 30-Jul-1998] (Gish, Warren (1994-1997). Unpublished; Gish, Warren and David J. States (1993). Identification of protein coding regions by database similarity search. Nat. Genet. 3:266-72).

The obtained sequences were functional genes which encoded intact and functional polypeptides, because they were obtained as ampicillin resistant clones as explained *supra*.

Example 2. Determining function by homology

The function of the polypeptides SEQ ID NO: 26 to SEQ ID NO: 50 were annotated by sequences comparison with genes or polypeptides of known function. The polypeptides of the invention were compared to a list of closest related sequences from public and inhouse databases of contig's. The contigs, from which SEQ ID NO: 26 to SEQ ID NO: 50 were derived, were subsequently compared to sequences available in standard public DNA and protein sequences databases by using the program BLASTX 2.0a19MP-WashU [14-Jul-1998]. A careful analysis of sequence alignments of SEQ ID NO: 26 to SEQ ID NO: 40 to their closest related sequences with known function from other databases made it possible to predict the function of these polypeptides on the basis of the degree of amino acid identity. Even when the overall amino acid identity was below 40%, which usually makes it difficult to make a good prediction, we were able to predict the function of SEQ ID NO: 26 to SEQ ID NO: 40 by carefully analysing and interpreting the amino acid residues in the catalytic sites or in important regions of the polypeptide sequences. If the amino acids of the catalytic site of a known sequences were also present in the polypeptide of the invention, combined with a sufficient overall amino acid identity, it was concluded that the polypeptide from Alicyclobacillus sp DSM 15716 had the same function as the known sequence.

Example 3 Preparing polypeptides of SEQ ID NO: 26 to SEQ ID NO: 50

To prepare the polypeptides of SEQ ID NO: 26 to SEQ ID NO: 50, the genes comprised in SEQ ID NO: 1 to SEQ ID NO: 25 encoding these polypeptides are expressed by fusing the DNA encoding the open reading frame to DNA a promoter, ribosome- binding site and terminator suitable for genes expression in an appropriate host strain, for example *Escherichia coli, Bacillus subtilis, Bacillus licheniformis* or *Bacillus clausii* or a derivative of *Alicyclobacillus sp.* The promoter can either be an inducible promotor or a constitutive promoter. Any signal sequences of SEQ ID NO: 26 to SEQ ID NO: 50 can be exchanged with a suitable signal peptide of another bacterium. The expression construct can either be part of a plasmid or of a linear DNA. It can be integrated into the chromosome of the host strain by recombination or it can be present in the host cell on a plasmid. Then the transformed cells carrying the gene of interest are grown in a suitable medium in the desired volume. If an inducible promoter is used, the gene expression is started by adding the inducer. Otherwise a no inducer is needed and the cells will be grown until a suitable amount of protein from the gene of interest is produced. Then the culture is harvested and the proteins are recovered by standard methods.

Example 4. Determining serine-carboxyl protease activity

The culture fluid or a cell lysate of a host strain synthesising and secreting a serine-carboxy protease in a suitable buffer may be assayed for that activity. A suitable volume of such a sample is spotted on agarose plates which contain the insoluble chromogenic substrate AZCL-collagen (Megazyme ™) or Azocoll (Sigma-Aldrich) and a suitable buffer at acidic pH, e.g. pH is 3-5. The plate is incubated for an appropriate time, e.g. one day at an appropriate temperature, e.g. 55□C. The activity is visible as blue halos around the spots. As an alternative to AZCL-collagen or Azocoll, non-labelled collagen is added to agar plates, on which enzyme activity can be detected as clearing zones. By addition of pepstatin, the protease activity of a serine carboxyl protease cannot be inhibited. As an alternative, the activity determination of a sample containing a serine-carboxyl protease can be measured as described in Tsuruoka N, Nakayama T, Ashida M, Hemmi H, Nakao M, Minakata H, Oyama H, Oda K, Nishino T; *"Collagenolytic serine-carboxyl proteinase from Alicyclobacillus sendaiensis strain NTAP-1: purification, characterization, gene cloning, and heterologous expression."* Appl Environ Microbiol. Vol. 69(1); pp 162-169; 2003 Jan.

Example 5. Determining multi-copper oxidase activity

The culture fluid or a cell lysate of a host strain synthesising and secreting a multi-copper oxidase in a suitable buffer may be assayed for that activity as described in Schneider et al., Enzyme and Microbial Technology 25, (1999) p. 502-508).

For example a suitable volume, which can be 15 microliter, of such a sample is spotted on agarose plates which contain ABTS (2,2'-Azinobis 3-Ethylbenzthiazolin-6-sulfonic acid) at a suitable concentration, e.g. 1mM, in a suitable puffer, e.g. 0,1 M sodium acetat buffer for pH 5,5. The plate is incubated for an appropriate time e.g. 16 hours, at an appropriate temperature, e.g. 55 °C. The activity is visible as a green zone around the sample. The assay works on supernatants and extracts.

Example 6. Determining serine protease activity

The culture fluid or a cell lysate of a host strain synthesising and secreting a serine protease in a suitable buffer may be assayed for that activity. A suitable volume of such a sample is spotted on agarose plates which contain the insoluble chromogenic substrate AZCL-casein (Megazyme ™) or AZCL-collagen (Megazyme ™) and a suitable buffer at suitable pH. The plate is incubated for an appropriate time, e.g. one day, at an appropriate temperature, e.g. 55□C. The ac-

tivity is visible as blue halos around the spots. As an alternative to AZCL-casein and AZCL-collagen (Megazyme ™) non-labelled casein or non-labelled collagene can be used. On non-labelled collagen or non-labelled casein spotted on agarose plates, clearing zones form in the presence of a serine protease.

Example 7. Determining aspartyl protease activity

The culture fluid or a cell lysate of a host strain synthesising and secreting an aspartyl protease in a suitable buffer was assayed for that activity. A suitable volume of such a sample can be spotted on agarose plates, which contain the insoluble chromogenic substrate AZCL-collagen (Megazyme ™) and a suitable buffer at acidic pH, e.g. pH is 3-5. The plate can be incubated for an appropriate time, e.g. one day, at an appropriate temperature, e.g. 55□C. The activity is visible as blue halos around the spots. As an alternative to AZCL-collagen, non-labelled collagen can be used. On non-labelled collagen spotted on agarose plates, clearing zones form in the presence of an aspartyl protease. Upon specifically testing the aspartyl protease of ID NO: 27; the activity was determined as a spot test of 20 microliter culture fluid on 0.1% AZCL-collagen (Megazyme ™) spotted on LB-PG agar plates at pH 3.4. The plates were incubated at 55 °C (over night) and the presence of the aspartyl protease was visible as blue halos around the spots.

The aspartyl protease comprised in SEQ ID NO: 27 showed significant sequence similarity to aspartyl peptidases belonging to family A4. This family contains peptidase sequences, which have an aspartate and a glutamate residue in their active site. Both residues were conserved in the aspartyl protease comprised in SEQ ID NO: 27. The aspartyl protease comprised in SEQ ID NO: 27 is thus the first bacterial polypeptide showing significant sequence similarity especially at the active sites with sequences from peptidase family A4 and therefore also the first bacterial A4 protease.

SEQ ID NO: 27 was compared to a reference sequence of family A4 peptidases; Aspergillus niger aspergillopepsin II (SEQ ID NO: 55; Swissprot P24665; Takahashi,K.; Inoue,H.; Sakai,K.; Kohama,T.; Kitahara,S.; Takishima,K.; Tanji,M.; Athauda,S.B.P.; Takahashi,T.; Akanuma,H.; Mamiya,G.; Yamasaki, M); *The primary structure of Aspergillus niger acid proteinase A.*; J. Biol. Chem.; Vol 266; p. 19480; 1991). This polypeptide contained a signal peptide (aa1-, and two propeptides (aa 19-58 and aa 99-109), which are removed after secretion during maturation. During maturation a heavy and a light chain are formed, which are cross-linked by disulfide bridges between cysteine residues. (Inoue,H.; Kimura,T.; Makabe,O.; Takahashi,K.; *The gene and deduced protein sequences of the zymogene of Aspergillus niger acid proteinase A*; J. Biol.

Chem.; vol. 266; p. 19484; 1991). The amino acids similar to the second propeptide (aa99-109) and the amino acids corresponding to the cross-linking cysteine residues of SEQ ID NO.55 are missing in SEQ ID 27 (see alignment). Only a fungal A4 peptidase has previously been described to to lack cysteine residues (Maita,T.; Nagata,S.; Matsuda,G.; Maruta,S.; Oda,K.; Murao,S.; Tsuru,D.; *Complete amino acid sequence of Scytalidium lignicolum acid protease B*; J. Biochem.; vol. 95; p. 465; 1984).

Alignment of SEQ ID NO: 55 with SEQ ID NO: 27

```
SWISSPROT_P24665   MKFSTILTGS-LFATAALAAPLTEKRRARKEARAAGKRHSNPPYIPGSDKEILKLNGTTN
Seq ID No.27       MNGTSVWKASGIAAASCLTAAALLAWPHATSTLDASPAIFHAPRHALSPNTSPKPNSVQA

                                                    ¤¤¤¤¤¤¤¤¤¤¤¤      :
SWISSPROT_P24665   EEY---SSNWAGAVLIGDGYTKVTGEFTVPSVSAGSSGSSGYGGGYGYWKNKRQSEEYCA
Seq ID No.27       QNFGWSASNWSGYAVTGSTYNDITGSWIVPAVSP----------------SKR--STYS-

                         *   :           :
SWISSPROT_P24665   SAWVGIDGDTCETAILQTGVDFCYEDGQTSYDAWYEWYPDYAYDFSDITISEGDSIKVTV
Seq ID No.27       SSWIGIDG-FNNSDLIQTGTEQDYVNGHAQYDAWWEILPAPETVISNMTIAPGDRMSAHI

                                                  :         *
SWISSPROT_P24665   EATSKSSGSATVENLTTGQSVTHTFSGNVEGDLCETNAEWIVEDFESGDSLVAFADFGSV
Seq ID No.27       HNNGNGTWTITLTDVTRNETFSTTQSYSGPG----SSAEWIQEAPEIGGRIATLANYGET

SWISSPROT_P24665   TFTNAEATSG--GSTVGPSDAT---------------------------------------
Seq ID No.27       TFDPGTVNGGNPGFTLVPTRATWCRTTRSCLCRPHPTRIPTASTWPTAPTSRAHRPPDPR

SWISSPROT_P24665   -----VMDIEQDGSVLTETSVSGDSVTVTYV------------
Seq ID No.27       RSRRPCMEAQGPASFFARTLAPSRDVAAHAPQGHRPSALVRRA
```

* = amino acids forming the active site in Swissprot P24665

: = cysteine residues froming disulfide bonds in Swissprot P24665

¤ = propeptide removed from the Swissprot P24665 zymogene.

Example 8. Determining acid beta-glucanase activity

The culture fluid or a cell lysate of a host strain synthesising and secreting an beta-glucanase in a suitable buffer may be assayed for that activity. A suitable volume of such a sample is spotted on agarose plates which contain the insoluble chromogenic substrate AZCL-beta-glucan (Megazyme ™) and a suitable buffer at acidic pH, e.g. pH is 3-5. The plate is incubated for an appropriate time, e.g. one day, at an appropriate temperature, e.g. 55□C. The activity is visible as blue halos around the spots.

Example 9. Determining acid phosphatase activity

A suitable volume of the culture fluid or a cell lysate of a host strain synthesising and secreting the acid phosphatase in a suitable buffer at a suitable pH at an appropriate temperature, e.g. 55 °C is incubated with para-nitrophenolphosphate (pNPP) for measuring the enzyme activity. The products of the enzymatic reaction are p-nitrophenol and inorganic phosphate or Pi. NaOH is added to end the phosphatase assay after a suitable reaction time and forms p-nitrophenolate. The absorbation of p-nitrophenolate is measured optically at 405 nm.

As an alternative, a suitable volume of the culture fluid or a cell lysate of a host strain synthesising and secreting the acid phosphatase in a suitable buffer at a suitable pH at an appropriate temperature, e.g. 55 °C is used for measuring the enzyme activity with the EnzChek™ Acid Phosphatase Assay Kit (E-12020) (Molecular Probes Europe BV; PoortGebouw, Rijnsburgerweg 10; 2333 AA Leiden, The Netherlands).

Example 10. Determining polysaccharide deacetylase activity

A suitable volume of the culture fluid or a cell lysate of a host strain synthesising and secreting the polysaccharide deacteylase in a suitable buffer at an appropriate temperature, e.g. 55□C is used for measuring the activity. Bacterial murein, *N,N'*-diacetylchitobiose (Sigma) or galactose pentaacetate (Sigma) or and cellulose acetate (Sigma) can be used as substrate(s) for this enzyme type. The acetate released from the substrate by the enzyme can be measured with an acetic acid assay kit (Biopharm) adapted for the physical requierments of the enzyme (Kosugi A, Murashima K, and Doi RH; *Xylanase and Acetyl Xylan Esterase Activities of XynA, a Key Subunit of the Clostridium cellulovorans Cellulosome for Xylan Degradation*; Appl. Environm. I Microbiol.; vol. 68; pp. 6399-6402; 2002)

Example 11. Determining endo-beta-N-acetylglucosaminidase activity

A suitable volume of the culture fluid or a cell lysate of a host strain synthesising and secreting the endo-beta-N-acetylglucosaminidase activity in a suitable buffer, e.g. pH 3-5, at an appropriate temperature, e.g. 55 °C can be used for measuring the activity in accordance with MH Rashid, M Mori and J Sekiguchi; *Glucosaminidase of Bacillus subtilis: cloning, regulation, primary structure and biochemical characterization*; Microbiology; vol. 141; pp. 2391-2404; 1995.

Example 12. Determining peptidyl proly-isomerase activity

A suitable volume of the culture fluid or a cell lysate of a host strain synthesising and secreting the polysaccharide deacteylase in a suitable buffer at an appropriate temperature, e.g. 55 °C is

used for measuring the activity. The activity can be determined in accordance to Fischer, G., Bang, H. and Mech, C.; *Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides.*; Biomed. Biochim. Acta; vol. 43; pp. 1101-1111;1984. This assay may be modified appropriately to suit the specific peptidyl proly-isomerase such as that comprised in SEQ ID NO: 36.

Example 13. Determining acid cellulase activity

The culture fluid or a cell lysate of a host strain synthesising and secreting an acid cellulase in a suitable buffer may be assayed for that activity. A suitable volume of such a sample is spotted on agarose plates which contain the insoluble chromogenic substrate AZCL-HE-cellulose (Megazyme ™) and a suitable buffer at acidic pH, e.g. pH is 3-5. The plate is incubated for an appropriate time, e.g. one day, at an appropriate temperature, e.g. 55□C. Presence of acid cellulase is visible as blue halos around the spots.

Example 14. Determining xylan deacetylase activity

A suitable volume of the culture fluid or a cell lysate of a host strain synthesising and secreting the polysaccharide deacteylase in a suitable buffer at an appropriate temperature, e.g. 55□C can be used for measuring xylan deacetylase activity. Xylan deacetylase activity is measured as acetate release from acetylated xylan, which is prepared from birchwood xylan by the method of Johnson et al. 1988 (Johnson, K. G., J. D. Fontana, and C. R. Mackenzie. 1988. Measurement of acetylxylan esterase in Streptomyces. Methods Enzymol. 160:551-560). The acetate released from acetyl xylan is measured with an acetic acid assay kit (Biopharm) adapted for the physical requierments of the enzyme (Kosugi A, Murashima K, and Doi RH; *Xylanase and Acetyl Xylan Esterase Activities of XynA, a Key Subunit of the Clostridium cellulovorans Cellulosome for Xylan Degradation*; Appl. Environm. I Microbiol.; vol. 68; pp. 6399-6402; 2002).

Example 15. Determining phytase activity

The culture fluid or a cell lysate of a host strain synthesising and secreting a phytase in a suitable buffer may be assayed for phytase activity. A suitable volume of such a sample is diluted in 0.1 M sodium acetate and 0.01% Tween-20, pH 5.5 in a suitable buffer, which can be -HCl at pH 3.0 to 3.5, sodium acetate at pH 4.0 to 5.5, morpholincethanesulfonic acid (MES) at pH 6.0 to 6.5, and Tris-HCl at pH 7.0 to 9.0, are further diluted in 26-fold into the substrate solution (5 mM sodium phytate [Sigma] in 0.1 M sodium acetate, and 0.01% Tween-20 [pH 5.5], and preincubated at 37°C) to start the reaction. After 30 min at 37°C, the reaction is stopped by adding an

equal volume of 10% trichloroacetic acid. Free inorganic phosphate is measured by the addition of an equal volume of molybdate reagent containing, in 100 ml, 7.3 g of $FeSO_4$, 1.0 g of $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, and 3.2 ml of H_2SO_4. Absorbance was measured at 750 nm (Vmax micro-titer plate reader; Molecular Devices) (Lassen SF; Breinholt J; Ostergaard PR; Brugger R; Bischoff A; Wyss M; Fuglsang CC; *Expression, gene cloning, and characterization of five novel phytases from four basidiomycete fungi: Peniophora lycii, Agrocybe pediades, a Ceriporia sp., and Trametes pubescens*; Appl. Environ. Micr.; 67; pp. 4701-4707; 2001).

Example 16. Determining phospholipase activity

The culture fluid or a cell lysate of a host strain synthesising and secreting a phospholipase in a suitable buffer may be assayed for phospholipase activity. Lecithin is added to suitable volume of such a sample. The Lecithin is hydrolyzed under constant pH and temperature, and the phospholipase activity is determined as the rate of titrant (0.1N NaOH) consumption during neutralization of the liberated fatty acid. The substrate is soy lecithin (L-α-Phosphotidyl-Choline), and the conditions are pH 8.00, 40.0°C, reaction time 2 min. The unit (LEU) is defined relative to a standard.

Example 17: Expression of aspartyl protease gene (SEQ ID NO: 2) in *Bacillus subtilis*.

The signal peptide from the protease SAVINASE™ (also known as subtilisin 309 from B. Licheniformis from Novozymes A/S) was fused by PCR in frame to the gene encoding the aspartyl protease (SEQ ID NO: **2**). The DNA coding for the resulting coding sequence was integrated by homologous recombination on the *Bacillus subtilis* host cell genome. The gene construct was expressed under the control of a triple promoter system (as described in WO 99/43835), consisting of the promoters from *Bacillus licheniformis* alpha-amylase gene (*amyL*), Bacillus amyloliquefaciens alpha-amylase gene (*amyQ*), and the *Bacillus thuringiensis cryIIIA* promoter including stabilizing sequence. The gene coding for Chloramphenicol acetyl-transferase was used as maker. (Described e.g in Diderichsen et al., *A useful cloning vector for Bacillus subtilis*. Plasmid, 30, p. 312, 1993).

Chloramphenicol resistant transformants were analyzed by DNA sequencing to veriify the correct DNA sequence of the construct. One such clone was selected.

Fermentations of the aspartyl protease (SEQ ID NO: **2**) expression clone was performed on a rotary shaking table in 500 ml baffled Erlenmeyer flasks each containing 100 ml PS-1 media supplemented with 6 mg/l chloramphenicol. The clone was fermented for 6 days at 37 °C and sample was taken at day 3, 4, 5 and 6 and analyzed for proteolytic activity. The

activity was determined (se example 7) as a spot test of 20 microliter culture fluid on 0.1% AZCL-collagen (Megazyme ™) LB-PG agar plates at pH 3.4. The plates were incubated at 55 °C (over night) and the activity was visible as blue halos around the spots.

Example 18: Purification and characterization of the family A4 protease from *Alicyclobacillus sp.*

Purification

Culture broth was centrifuged (20000 x g, 20 min) and the supernatants were carefully decanted from the precipitates. The combined supernatants were filtered through a Seitz EKS plate in order to remove the rest of the *Bacillus* host cells. The EKS filtrate was adjusted to pH 4.0 with citric acid and heated to 70°C with good stirring on a water bath. When the solution reached 70°C (it took approx. 15 minutes to get from 25°C to 70°C), the solution was immediately placed on ice. This heat treatment resulted in some precipitation, which was removed by another Seitz EKS filter plate filtration. Ammonium sulfate was added to the second EKS fitrate to 1.6M final concentration and the pool was applied to a Butyl Toyopearl S column equilibrated in 20mM CH_3COOH/NaOH, 1.6M $(NH_4)_2SO_4$, pH 4.5. After washing the Butyl column extensively with the equilibration buffer, the enzyme was eluted with a linear $(NH_4)_2SO_4$ gradient (1.6 --> 0M) in the same buffer. Fractions from the column were analysed for protease activity (using the pH 4.0 Assay buffer and 37°C assay temperature) and fractions with activity were pooled. The pooled fractions were transferred to 20mM CH_3COOH/NaOH, pH 5.5 on a G25 sephadex column and applied to a SOURCE 30Q column equilibrated in the same buffer. After washing the SOURCE 30Q column extensively with the equilibration buffer, the protease was eluted with a linear NaCl gradient (0 --> 0.5M) in the same buffer. Fractions from the column were analysed for protease activity (pH 4.0, 37°C) and fractions with activity were pooled. The pool, which was slightly coloured, was treated with 1% (w/v) Activated charcoal for 5 minutes and the charcoal was removed by a 0.45µ filtration. The purity of the filtrate was analysed by SDS-PAGE, where only one band was seen on the coomassie stained gel.

Assay:

A Protazyme OL (cross-linked and dyed collagen) assay was used. A Protazyme OL tablet (from Megazyme) was suspended in 2.0ml 0.01% Triton X-100 by gentle stirring. 500 microliter of this suspension and 500 microliter assay buffer were mixed in an Eppendorf tube and placed on ice. 20 microliter protease sample (diluted in 0.01% Triton X-100) was added. The assay was initiated by transferring the Eppendorf tube to an Eppendorf thermomixer, which was set to the assay temperature. The tube was incubated for 15 minutes on the Eppendorf thermomixer at its highest shaking rate (1400 rpm). The incubation was stopped by transferring the tube back to the ice bath. Then the tube was centrifuged in an ice-cold centrifuge for a few minutes, 200 microliter supernatant was transferred to a microtiter plate and OD_{650} was read at 650 nm. A buffer blind was included in the assay (instead of enzyme). OD_{650}(enzyme) – OD_{650}(buffer blind) was a measure of protease activity."

Protease assay:

Substrate : Protazyme OL tablets (Megazyme T-PROL).

Temperature : Controlled.

Assay buffers : 100mM succinic acid, 100mM HEPES, 100mM CHES, 100mM CABS, 1mM $CaCl_2$, 150mM KCl, 0.01% Triton X-100 adjusted to pH-values 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0 and 12.0 with HCl or NaOH.

Characterisation: pH-activity , pH-stability, and temperature-activity:

The above protease assay was used for obtaining the pH-activity profile, the pH-stability profile as well as the temperature-activity profile at pH 3.0. For the pH-stability profile the protease was diluted 5x in the Assay buffers and incubated for 2 hours at 37°C. After incubation the protease samples were transferred to pH 3.0, before assay for residual activity, by dilution in the pH 3 Assay buffer.

pH-activity profile at 37°C

pH	Alicyclobacillus protease from EXP00663
2	0.90
3	0.98
4	1.00
5	0.93
6	0.77
7	0.28
8	0.04
9	0.02

pH-stability profile (residual activity after 2 hours at 37°C)

pH	Alicyclobacillus protease from EXP00663
2.0	0.93
3.0	0.97
4.0	0.94

5.0	0.97
6.0	0.93
7.0	0.94
8.0	0.99
9.0	0.94
10.0	0.81
11.0	0.76
12.0	0.46
3.0 and after 2 hours at 5 °C	1.00

Temperature activity profile (at pH 3.0)

Temp (°C)	Alicyclobacillus protease from EXP00663
15	0.08
25	0.19
37	0.60
50	0.94
60	1.00
70	0.89
80	0.45

Other characteristics:

The relative molecular weight of the A4 protease as determined by SDS-PAGE was:
M_r = 26 kDa.

Example 19: Expression of acid cellulase gene (SEQ ID NO: 1) in *Bacillus subtilis*.

The signal peptide from Termamyl™ (Novozymes) was fused by PCR in frame to the gene encoding the acid cellulase (SEQ ID NO: 1). The DNA coding for the resulting coding sequence was integrated by homologous recombination on the *Bacillus subtilis* host cell genome. The gene construct was expressed under the control of a triple promoter system (as described in WO 99/43835), consisting of the promoters from *Bacillus licheniformis* alpha-amylase gene (*amyL*), Bacillus amyloliquefaciens alpha-amylase gene (*amyQ*), and the *Bacillus thuringiensis cryIIIA* promoter including stabilizing sequence. The gene coding for Chloramphenicol acetyl-transferase was used as maker (Described e.g in Diderichsen et al., *A useful cloning vector for Bacillus subtilis*. Plasmid, 30, p. 312, 1993).

Chloramphenicol resistant transformants were analyzed by DNA sequencing to verify the correct DNA sequence of the construct. One such clone was selected.

Fermentations of the acid cellulase (SEQ ID NO: 1) expression clone was performed on a rotary shaking table in 500 ml baffled Erlenmeyer flasks each containing 100 ml PS-1 media supplemented with 6 mg/l chloramphenicol. The clone was fermented for 3 days at 37 °C and sample was taken at day 1,2 and 3 and analyzed for cellulase activity. The activity was determined as a spot test of 20 microliter culture fluid on 0.1% AZCL-HE-cellulase (Megazyme ™) LB-PG agar plates at pH 3.4. The plates were incubated at 55 °C (over night) and the activity was visible as blue halos around the spots.